Ecological Assessment of Streams in the Coal Mining Region of West Virginia Using Data Collected by the U.S. EPA and Environmental Consulting Firms

February 2003

Prepared by:

Florence Fulk and Bradley Autrey
U.S. Environmental Protection Agency
National Exposure Research Laboratory
Cincinnati, Ohio

John Hutchens
Coastal Carolina University
Conway, South Carolina

Jeroen Gerritsen, June Burton, Catherine Cresswell, and Ben Jessup
Tetra Tech, Inc.
Owings Mills, Maryland

U.S. Environmental Protection Agency
National Exposure Research Laboratory
26 W. Martin Luther King Drive
Cincinnati, Oh 45268

NOTICE

This research described in this report has been funded wholly or in part by the U.S. Environmental Protection Agency. This document has been prepared at the U.S. Environmental Protection Agency, National Exposure Research Laboratory, Ecological Exposure Research Division in Cincinnati, Ohio.

EXECUTIVE SUMMARY

INTRODUCTION

Recently, the Mountaintop Mining (MTM) and Valley Fill (VF) operations in the Appalachian Coal Region have increased. In these operations, the tops of mountains are removed, coal materials are mined and the excess materials are deposited into adjacent valleys and stream corridors. The increased number of MTM/VF operations in this region has made it necessary for regulatory agencies to examine the relevant regulations, policies, procedures and guidance needed to ensure that the potential individual and cumulative impacts are considered. This necessity has resulted in the preparation of an Environmental Impact Statement (EIS) concerning the MTM/VF activities in West Virginia. The U.S. Environmental Protection Agency (EPA), U.S. Army Corps of Engineers, U.S. Office of Surface Mining, and U.S. Fish and Wildlife Service, in cooperation with the West Virginia Department of Environmental Protection, are working to prepare the EIS. The purpose of the EIS is to establish an information foundation for the development of policies, guidance and coordinated agency decision-making processes to minimize, to the greatest practicable extent, the adverse environmental effects to the waters, fish and wildlife resources in the U.S. from MTM operations, and to other environmental resources that could be affected by the size and location of fill material in VF sites. Furthermore, the EIS's purpose is to determine the proposed action, and develop and evaluate a range of reasonable alternatives to the proposed action.

The U.S. EPA's Region 3 initiated an aquatic impacts study to support the EIS. From the spring 1999 through the winter 2000, U.S. EPA Region 3 personnel facilitated collection of water chemistry, habitat, macroinvertebrate and fish data from streams within the MTM/VF Region. In addition, data were also collected by three environmental consulting firms, representing four coal mining companies. The National Exposure Research Laboratory (NERL) of the U.S. EPA's Office of Research and Development assembled a database of U.S. EPA and environmental consulting firm data collected from the MTM/VF Region. Using this combined data set, NERL analyzed fish and macroinvertebrate data independently to address two study objectives: 1) determine if the biological condition of streams in areas with MTM/VF operations is degraded relative to the condition of streams in unmined areas and 2) determine if there are additive biological impacts to streams where multiple valley fills are located. The results of these analyses, regarding the aquatic impacts of MTM/VF operations, are provided in this report for inclusion in the overall EIS.

ANALYTICAL APPROACH AND RESULTS

Fish Data Analyses and Results

The Mid-Atlantic Highlands Index of Biotic Integrity (IBI), was used in the analyses of the fish data. This index is made up of scores from multiple metrics that are responsive to stress. Each of the sites sampled was placed into one of six EIS classes (i.e., Unmined, Filled, Mined, Filled/Residential, Mined/Residential, Additive). Due to inadequate sample size, the Mined/Residential class was removed from analyses. The Additive class was analyzed separately because it was made up of sites that were potentially influenced by multiple sources of stress.

The objective of the IBI analyses were to examine and compare EIS classes to determine if they are associated with the biological condition of streams. The distributions of IBI scores showed that the Filled and Mined classes had lower overall IBI scores than the other EIS classes. The Filled/Residential class had higher IBI scores than the Filled or Mined classes. The combined Filled/Residential class and the Unmined class had median scores that were similar to regional reference sites. Unmined and regional reference sites were primarily in the "fair" range and a majority of the Filled/Residential sites fell within the "good" range.

A standard Analysis of Variance (ANOVA) was used to test for differences among EIS classes and the Least Square (LS) Means procedure using Dunnett's adjustment for multiple comparisons tested whether the Filled, Filled/Residential, and Mined EIS classes were significantly different ($p < 0.01$) from the Unmined class. The ANOVA showed that there were significant differences among EIS classes. The LS Means test showed that the IBI scores from Filled and Mined sites were significantly lower than the IBI scores from Unmined sites, and the IBI scores from Filled/ Residential sites were significantly higher than the IBI scores from Unmined sites. Of the nine metrics in the IBI, only the Number of Minnow Species and the Number of Benthic Invertivore Species were significantly different in the Unmined class. Therefore, it was determined that the primary causes of reduced IBI scores in Filled and Mined sites were the reductions in these two metrics relative to the Unmined sites.

It was found that Filled, Mined, and Filled/Residential sites in watersheds with areas greater than 10 km^2 had "fair" to "good" IBI scores, while Filled and Mined sites in watersheds with areas less than 10 km^2 often had "poor" IBI scores. Of the 14 sites Filled and Mined) in watersheds with areas greater than 10 km^2, four were rated "fair" and ten were rated "good" or better. Of the 17 sites (Filled and Mined) in watersheds with areas less than 10 km^2, only three were rated "fair" and 14 were rated "poor". The effects of fills were statistically stronger in watersheds with areas less than 10 km^2. Filled sites had IBI scores that were an average of 14 points lower than Unmined sites. It is possible that the larger watersheds act to buffer the effects of stress.

Additive sites were considered to be subject to multiple, and possibly cumulative, sources, and were not included in the analysis of the EIS classes reported above. From the additive analysis, it was determined that the Twelvepole Creek Watershed, in which the land use was mixed residential and mining, had "fair" IBI scores in most samples, and there are no apparent additive effects of the land uses in the downstream reaches of the watershed. Also, Twentymile Creek, which has only mining-related land uses, may experience impacts from the Peachorchard tributary. The IBI scores appear to decrease immediately downstream of the confluence of the two creeks, whereas above the confluence, IBI scores in the Twentymile Creek are higher than in the Peachorchard Creek. Peachorchard Creek may contribute contaminants or sediments to Twentymile Creek, causing degradation of the Twentymile IBI scores downstream of Peachorchard Creek.

The correlations between IBI scores and potential stressors detectable in water were examined. Zinc, sodium, nickel, chromium, sulfate, and total dissolved solids were associated with reduced IBI scores. However, these correlations do not imply causal relationships between the water quality parameters and fish community condition.

Macroinvertebrate Data Analyses and Results

The benthic macroinvertebrate data were analyzed for statistical differences among EIS classes. Macroinvertebrate data were described using the WVSCI and its component metrics. The richness metrics and the WVSCI were rarefied to 100 organisms to adjust for sampling effort. Four EIS classes (i.e.; Unmined, Filled, Mined, and Filled/Residential) were compared using one-way ANOVAs. Significant differences among EIS classes were followed by the Least Square (LS) Means procedure using Dunnett's adjustment for multiple comparisons to test whether the Filled, Filled/Residential, and Mined EIS classes were significantly different ($p < 0.01$) from the Unmined class. Comparisons were made for each of the sampling seasons where there were sufficient numbers of samples.

The results of the macroinvertebrate analyses showed significant differences among EIS classes for the WVSCI and some of its component metrics in all seasons except autumn 2000. Differences in the WVSCI were primarily due to lower Total Taxa, especially for mayflies, stoneflies, and caddisflies, in the Filled and Filled/Residential EIS classes. Sites in the Filled/Residential EIS class usually scored the worst of all EIS classes across all seasons.

Using the mean values for water chemistry parameters at each site, the relationships between WVSCI scores and water quality were determined. The strongest of these relationships were negative correlations between the WVSCI and measures of individual and combined ions. The WVSCI was also negatively correlated with the concentrations of Beryllium, Selenium, and Zinc.

Multiple sites on the mainstem of Twentymile Creek were identified as Additive sites and were included in an analysis to evaluate impacts of increased mining activities in the watershed across seasons and from upstream to downstream of the Twentymile Creek. Sites were sampled during four seasons. Pearson correlations between cumulative river kilometer and the WVSCI and it's component metrics were calculated. The number of metrics that showed significant correlations with distance along the mainstem increased across seasons. The WVSCI was significantly correlated with cumulative river kilometer in Winter 2000, Autumn 2000 and Winter 2001. For Winter 2001, a linear regression of the WVSCI with cumulative river kilometer indicated that the WVSCI decreased approximately one point upstream to downstream for every river kilometer.

MAJOR FINDINGS AND SIGNIFICANCE

Fish Data Findings and Significance

It was determined that IBI scores were significantly reduced at Filled sites compared to Unmined sites by an average of 10 points, indicating that fish communities were degraded below VFs. The IBI scores were similarly reduced at sites receiving drainage from historic mining or contour mining (i.e., Mined sites) compared to Unmined sites. Nearly all Filled and Mined sites with catchment areas smaller than 10 km^2 had “poor” IBI scores. At these sites, IBI scores from Filled sites were an average of 14 points lower than the IBI scores from Unmined sites. Filled and Mined sites with catchment areas larger than 10 km^2 had “fair” or “good” IBI scores. Most of the Filled/Residential sites were in these larger watersheds and tended to have “fair” or “good” IBI scores.

It was also determined that the Twelvepole Creek Watershed, which had a mix of residential and mining land uses, had “fair” IBI scores in most samples; there were no apparent additive effects of the land uses in the downstream reaches of the watershed. Twentymile Creek, which had only mining-related land uses, had “good” IBI scores upstream of its confluence with Peachorchard Creek, and “fair” and “poor” scores for several miles downstream of its confluence with Peachorchard Creek. Peachorchard Creek had “poor” IBI scores, and may have contributed to the degradation of the Twentymile Creek’s IBI scores downstream of their confluence.

Macroinvertebrate Data Findings and Significance

The macroinvertebrate analyses showed significant differences among EIS classes for the WVSCI and some of its metrics in all seasons except autumn 2000. Differences in the WVSCI were primarily due to lower Total Taxa and lower EPT Taxa in the Filled and Filled/Residential EIS classes. Sites in the Filled/Residential EIS class usually had the lowest scores of all EIS classes across all seasons. It was not determined why the Filled/Residential class scored worse than the Filled class alone. U.S. EPA (2001 Draft) found the highest concentrations of sodium in the Filled/Residential EIS class, which may have negatively impacted these sites compared to those in the Filled class.

When the results for Filled and Unmined sites alone were examined, significant differences were observed in all seasons except autumn 1999 and autumn 2000. The lack of differences between Unmined and Filled sites in autumn 1999 was due to a decrease in Total Taxa and EPT Taxa at Unmined sites relative to the summer 1999. These declines in taxa richness metrics in Unmined sites were likely the result of drought conditions. Despite the relatively drier conditions in Unmined sites during autumn 1999, WVSCI scores and EPT Taxa richness increased in later seasons to levels seen in the spring 1999, whereas values for Filled sites stayed relatively low.

In general, statistical differences between the Unmined and Filled EIS classes corresponded to ecological differences between classes based on mean WVSCI scores. Unmined sites scored "very good" in all seasons except autumn 1999 when the condition was scored as "good". The conditions at Filled sites ranged from "fair" to "good". However, Filled sites that scored "good" on average only represented conditions in the Twentymile Creek watershed in two seasons (i.e., autumn 2000 and winter 2001). These sites are not representative of the entire MTM/VF study area. On average, Filled sites had lower WVSCI scores than Unmined sites.

The consistently higher WVSCI scores and the Total Taxa in the Unmined sites relative to Filled sites across six seasons showed that Filled sites have lower biotic integrity than sites without VFs. Furthermore, reduced taxa richness in Filled sites is primarily the result of fewer pollution-sensitive EPT taxa. The lack of significant differences between these two EIS classes in autumn 1999 appears to be due to the effects of greatly reduced flow in Unmined sites during a severe drought. Continued sampling at Unmined and Filled sites would improve the understanding of whether MTM/VF activities are associated with seasonal variation in benthic macroinvertebrate metrics and base-flow hydrology.

Examination of the Additive sites from the mainstem of Twentymile Creek indicated that impacts to the benthic macroinvertebrate communities increased across seasons and upstream to downstream of Twentymile Creek. In the first sampling season one metric, Total Taxa, was negatively correlated with distance along the mainstem. The number of metrics showing a relationship with cumulative river mile increased across seasons, with four of the six metrics having significant correlations in the final sampling season, Winter 2001. Also in Winter of 2001, a regression of the WVSCI versus cumulative river kilometer estimates a decrease of approximately one point in the WVSCI for each river kilometer. Season and cumulative river kilometer in this dataset may be surrogates for increased mining activity in the watershed.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

Section **Page**

TABLE OF CONTENTS (CONTINUED)

TABLES

TABLES (CONTINUED)

Table **Page**

FIGURES

Figure **Page**

ACKNOWLEDGMENTS

This report could not have been completed without the efforts of many individuals and organizations. We would like to thank the U.S. EPA Region 3 personnel, especially Jim Green, Maggie Passemore, Frank Borsuk, Gary Bryant and Bill Hoffman for providing data, guidance and support for this study. We would like to thank Hope Childers of the Center for Educational Technologies at the Wheeling Jesuit University for her role in supporting the U.S. EPA Region 3 in this study. We would like to thank the Pennsylvania State University's School of Forest Resources, especially Jay Stauffer, Jr. and C. Paola Ferreri for providing data in support of this study and the U.S. Fish and Wildlife Service for supporting their work.

We would also like to thank Biological Monitoring, Incorporated; Potesta & Associates, Incorporated; and Research, Environmental, and Industrial Consultants, Incorporated for collecting data in support of this study. We also thank Arch Coal, the Massey Energy Company, the Penn Coal Corporation, the Fola Coal Company and the West Virginia Coal Association for providing access to sampling sites and supporting the collection of data.

We are grateful to Ken Fritz and David M. Walters of the U.S. EPA's National Exposure Research Laboratory and Lori Winters of ORISE for reviewing this document. We are also grateful to Alicia Shelton of SoBran Environmental for her efforts in editing and formatting this document.

1. INTRODUCTION

1.1. Background

Since the early 1990s, the nature and extent of coal mining operations in the Appalachian Region of the U.S. have changed. An increased number of large (> 1,200-ha) surface mines have been proposed and technology has allowed for the expanded role of Mountaintop Mining (MTM) and Valley Fill (VF) operations. In these operations, the tops of mountains are removed in order to make the underlying coal accessible (Figure 1-1). The excess materials from the mountaintop removals typically have been deposited into adjacent valleys and their stream corridors (Figure 1-2). These depositions cover perennial streams, wetlands and tracts of wildlife habitat. Given the increased number of mines and the increased scale of mining operations in the MTM/VF Region, it has become necessary for federal and state agencies to ensure that the relevant regulations, policies, procedures and guidance adequately consider the potential individual and cumulative impacts that may result from these projects (U.S. EPA 1999).

1.2. Environmental Impact Statement Development

The U.S. Environmental Protection Agency (EPA), U.S. Army Corps of Engineers (COE), U.S. Office of Surface Mining (OSM), and U.S. Fish and Wildlife Service (FWS), in cooperation with the West Virginia Department of Environmental Protection (DEP), are preparing an Environmental Impact Statement (EIS) concerning the MTM/VF activities in West Virginia. The purpose of developing the EIS is to facilitate the informed consideration of the development of policies, guidance and coordinated agency decision-making processes to minimize, to the greatest extent practicable, the adverse environmental effects to the waters, fish and wildlife resources in the U.S. from MTM operations, and to other environmental resources that could be affected by the size and location of fill material in VF sites (U.S. EPA 2001). Additionally, The EIS will determine the proposed action, and develop and evaluate a range of reasonable alternatives to the proposed action.

The goals of the EIS are to: (1) achieve the purposes stated above; (2) assess the mining practices currently being used in West Virginia; (3) assess the additive effects of MTM/VF operations; (4) clarify the alternatives to MTM; (5) make environmental evaluations of individual mining projects; (6) improve the capacity of mining operations, regulatory agencies, environmental groups and land owners to make informed decisions; and (7) design improved regulatory tools (U.S. EPA 2000). The major components of the EIS will include: human and

Figure 1-1. A MTM operation in West Virginia. The purpose of these operations are to remove mountaintops in order to make the underlying coal accessible.

Figure 1-2. A VF in operation. The excess materials from a MTM operation are being placed in this adjacent valley.

community impacts (i.e., quality of life, economic), terrestrial impacts (i.e., visuals, landscape, biota), aquatic impacts and miscellaneous impacts (i.e., blasting, mitigation, air quality).

1.3. Aquatic Impacts Portion of the EIS

The U.S. EPA's Region 3 initiated an aquatic impacts study to support the EIS. From the spring (i.e., April to June) 1999 through the winter (i.e., January to March) 2000, the U.S. EPA Region 3 collected data from streams within the MTM/VF Region. These data include water chemistry, habitat, and macroinvertebrates. With cooperation and guidance from the U.S. EPA Region 3, the Pennsylvania State University's (PSU's) School of Forest Resources collected fish data from streams in the MTM/VF Region. In addition to the data that were collected by the U.S. EPA Region 3 and PSU, data were also collected by three environmental consulting firms, representing four coal mining companies. These environmental consulting firms were Biological Monitoring, Incorporated (BMI); Potesta & Associates, Incorporated (POTESTA); and Research, Environmental, and Industrial Consultants, Incorporated (REIC).

Three reports which describe the data collected by the U.S. EPA Region 3 and PSU's School of Forest Resources were prepared. The first report summarized the condition of streams in the MTM/VF Region based on the macroinvertebrate data that were collected (Green et al. 2000 Draft). This report provided a descriptive analysis of the macroinvertebrate data. The second report described the fish populations in the MTM/VF Region based on the fish data collected by the PSU's School of Forest Resources (Stauffer and Ferreri 2000 Draft). This report used a fish index that was developed by the Ohio EPA for larger streams. The third report was a survey of the water quality of streams in the MTM/VF Region based on the water chemistry data collected by the U.S. EPA Region 3 (U.S. EPA 2002 Draft).

1.4. Scope and Objectives of This Report

In this document, the National Exposure Research Laboratory (NERL) of the U.S. EPA's Office of Research and Development (ORD) has assembled a database of Region 3, PSU and environmental consulting firm data collected from the MTM/VF Region. Using this combined data set, NERL analyzed fish and macroinvertebrate data separately to address the study's objectives. The results of these analyses will allow NERL to provide a report on the aquatic impacts of the MTM/VF operations for inclusion in the EIS.

The objectives of this document are to: 1) determine if the biological condition of streams in areas with MTM/VF operations is degraded relative to the condition of streams in unmined areas and 2) determine if there are additive biological impacts in streams where multiple VFs are located.

1.5. Biological Indices

One of the ways in which biological condition is assessed is through the use of biological indices. Biological indices allow stream communities to be compared by using their diversity, composition and functional organization. The use of biological indices is recommended by the Biological Criteria portion of the U.S. EPA's National Program Guidance for Surface Waters (U.S. EPA 1990). As of 1995, 42 states were using biological indices to assess impacts to streams (U.S. EPA 1996).

Two indices were identified as being appropriate for use with data collected from the MTM/VF Region. These were the Mid-Atlantic Highlands Index of Biotic Integrity (IBI) for fish (McCormick et al. 2001) and the West Virginia Stream Condition Index (WVSCI) for invertebrates (Gerritsen et al. 2000).

Due to the lack of a state developed fish index for West Virginia, an index created for use in the Mid-Atlantic Highlands was selected for evaluation of the fish data. The Mid-Atlantic Highlands IBI (McCormick et al. 2001) was developed using bioassessment data collected by the U.S. EPA from 309 wadeable streams from 1993 to 1996 in the Mid-Atlantic Highlands portion of the U.S. These data were collected using the U.S. EPA's Environmental Monitoring and Assessment Program (EMAP) protocols (Lazorchak et al. 1998). Site selection was randomly stratified. Fish were collected within reaches whose lengths were 40 times the wetted width of the stream with minimum and maximum reach lengths being 150 and 500 m, respectively. All fish collected for these bioassessments were identified to the species taxonomic level. An Analysis of Variance (ANOVA) showed that there were no differences between the ecoregions in which the data were collected. A subset of the data was used to develop the IBI and another subset was used to validate the IBI and its component metrics. Fifty-eight candidate metrics were evaluated. Of these, 13 were rejected because they did not demonstrate an adequate range, two were rejected because they had excessive signal-to-noise ratios, three were rejected because they were redundant with other metrics, one was rejected because it remained correlated with watershed area after it had been adjusted to compensate for area and 30 were rejected because they were not significantly correlated with anthropogenic impacts. The remaining nine metrics used in the IBI are described in Table 1-2 (McCormick et al. 2001). All metrics were scored on a continuous scale from 0 to 10. Three sets of reference condition criteria (i.e., least restrictive, moderately restrictive, most restrictive) were used to determine the threshold values for the metrics. For the metrics which decrease with perturbation (Table 1-1), a score of 0 was given if the value was less than the 5th percentile of the values from non-reference sites and a score of 10 was given if the value was greater than the 50th percentile of the values from reference sites defined by the most restrictive criteria. For the metrics which increase with perturbation (Table 1-1), a score of 0 was given if the value was greater than the 90th percentile of the values from non-reference sites and a score of 10 was given if the value was less than the 50th percentile of the values from reference sites defined by the moderately restrictive criteria. The IBI scores were scaled from 0 to 100 by summing the scores from the nine metrics and multiplying this sum by 1.11.

Table 1-1. The nine metrics in the Mid-Atlantic Highlands IBI, their definitions and their expected responses to perturbations.

Metric	Metric Description	Predicted Response to Stress
Native Intolerant Taxa	Number of indigenous taxa that are sensitive to pollution; adjusted for drainage area	Decrease
Native Cyprinidae Taxa	Number of indigenous taxa in the family Cyprinidae (carps and minnows); adjusted for drainage area	Decrease
Native Benthic Invertivores	Number of indigenous bottom dwelling taxa that consume invertebrates; adjusted for drainage area	Decrease
Percent Cottidae	Percent individuals of the family Cottidae (i.e., sculpins)	Decrease
Percent Gravel Spawners	Percent individuals that require clean gravel for reproductive success	Decrease
Percent Piscivore/Invertivores	Percent individuals that consume fish or invertebrates	Decrease
Percent Macro Omnivore	Percent individuals that are large and omnivorous	Increase
Percent Tolerant	Percent individuals that are tolerant of pollution	Increase
Percent Exotic	Percent individuals that are not indigenous	Increase

The WVSCI (Gerritsen et al. 2000) was developed using bioassessment data collected by the WVDEP from 720 sites in 1996 and 1997. These data were collected using the U.S. EPA's Rapid Bioassessment Protocols (RBP, Plafkin et al. 1989). From these bioassessments, 100 benthic macroinvertebrates were identified to the family taxonomic level from each sample. The information derived from the analyses of these data were used to establish appropriate site classifications for bioassessments, determine the seasonal differences among biological metrics, elucidate the appropriate metrics to be used in West Virginia and define the thresholds that indicate the degree of comparability of streams to a reference condition. The analyses of these data showed that there was no benefit to partitioning West Virginia into ecoregions for the purpose of bioassessment. The analyses also showed that variability in the data could be reduced by sampling only from late spring through early summer. Using water quality and habitat criteria, the reference and impaired sites were identified among the 720 sampled sites. Then, a suite of candidate metrics were evaluated based on their abilities to differentiate between reference and impaired sites, represent different aspects of the benthic macroinvertebrate community (i.e., composition, richness, tolerance), and minimize redundancy among individual component metrics. Based on these evaluations, it was determined that the metrics making up the WVSCI should be EPT taxa, Total taxa, % EPT, % Chironomidae, the Hilsenhoff Biotic Index (HBI) and % 2 Dominant taxa (Table 1-2). Next, the values for these metrics were calculated for all 720 sites and those values were standardized by converting them to a 0-to-100-point scale. The standardized scores for the six metrics were averaged for each site in order to obtain index scores. Data collected from West Virginia in 1998 were used to test the index. This analysis showed that the index was able to discriminate between reference and impaired sites (Gerritsen et al. 2000).

Table 1-2. The six metrics in the WVSCI, their definitions and their expected responses to perturbations.

Metric	Definition	Expected Response to Perturbation
EPT Taxa	The total number of EPT taxa.	Decrease
Total Taxa	The total number of taxa.	Decrease
% EPT	The percentage of the sample made up of EPT individuals.	Decrease
% Chironomidae	The percentage of the sample made up of Chironomidae individuals.	Increase
HBI	An index used to quantify an invertebrate assemblage's tolerance to organic pollution.	Increase
% 2 Dominant taxa	The percentage of the sample made up of the dominant two taxa in the sample.	Increase

2. METHODS AND MATERIALS

2.1. Data Collection

The U.S. EPA Region 3 collected benthic macroinvertebrate and habitat data from spring 1999 through spring 2000. These data were collected from 37 sites in five watersheds (i.e., Mud River, Spruce Fork, Clear Fork, Twentymile Creek, and Island Creek Watersheds) in the MTM/VF Region of West Virginia (Figure 2-1). Two sites were added to the study in spring 2000. These additions were a reference site not located near any mining activities and a supplementary site located near mining activities. Using these data, the U.S. EPA Region 3 developed a report (Green et al. 2000 Draft) which characterized the benthic macroinvertebrate assemblages in the MTM/VF Region of West Virginia.

The PSU's School of Forest Resources collected fish data in the MTM/VF Region of West Virginia and Kentucky. These data were collected from 58 sites in West Virginia and from 15 sites in Kentucky. The data collected from the Kentucky sites will not be used in this document. All of PSU's West Virginia sites were located in the same five watersheds from which the U.S. EPA Region 3 collected benthic macroinvertebrate, habitat and water quality data and most of these sites were located near the locations from which the U.S. EPA Region 3 collected these data. Data were collected in autumn 1999 and spring 2000. The results of this study were reported by Stauffer and Ferreri (2000 Draft).

The U.S. EPA Region 3 collected water quality data and water samples for chemical analyses from October 1999 through February 2001. These data were collected from the same 37 sites from which the U.S. EPA Region 3 collected benthic macroinvertebrate and habitat data. Using these data, the U.S. EPA Region 3 developed a report (U.S. EPA 2002 Draft) which characterized the water quality of streams in the MTM/VF Region of West Virginia.

The environmental consulting firm, BMI, collected water quality, water chemistry, habitat, benthic macroinvertebrate and fish data in the MTM/VF Region of West Virginia. These data were collected for Arch Coal, Incorporated from 37 sites in the Twentymile Creek Watershed and for Massey Energy Company from 11 sites in the Island Creek Watershed.

In addition, the environmental consulting firm, REIC, collected water quality, water chemistry, habitat, benthic macroinvertebrate and fish data in the MTM/VF Region of West Virginia. These data were collected for the Penn Coal Corporation from 18 sites in the Twelvepole Creek Watershed. Although the Twelvepole Creek Watershed is not among the

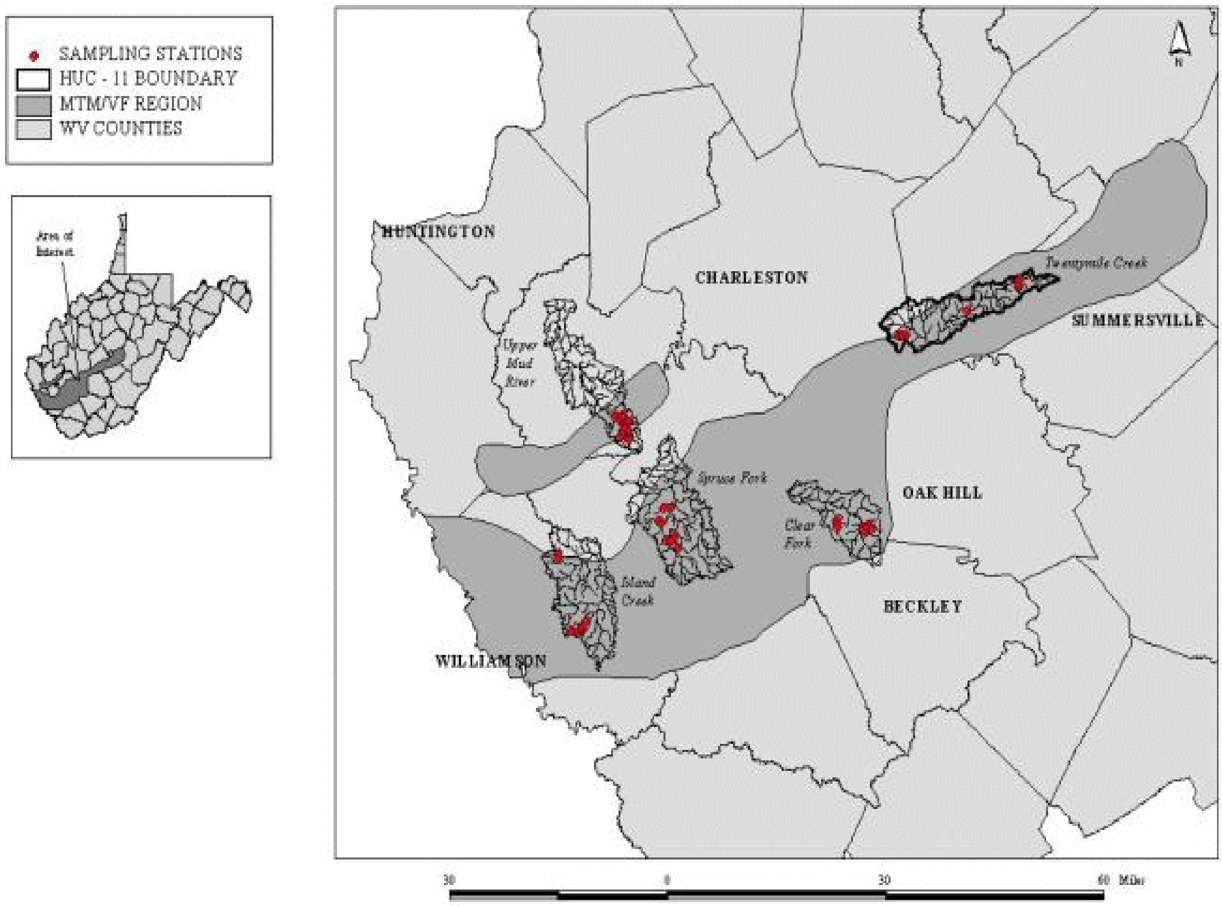


Figure 2-1. Study area for the aquatic impacts study of the MTM/VF Region of West Virginia.

watersheds from which the U.S. EPA Region 3 collected ecological data, some of these data will be considered in this report.

Finally, the environmental consulting firm, POTESTA, collected water quality, water chemistry, habitat, benthic macroinvertebrate, and fish data in the MTM/VF Region of West Virginia. These data were collected for the Fola Coal Company from ten sites in the Twentymile Creek Watershed (See Appendix E for a summary of benthic methods used by all groups).

2.2. Site Classes

Each of the sites sampled by the U.S. EPA Region 3, PSU or one of the participating environmental consulting firms was placed in one of six classes. These six classes were: 1)

Unmined, 2) Filled, 3) Mined, 4) Filled/Residential, 5) Mined/Residential and 6) Additive. The Unmined sites were located in areas where there had been no mining activities upstream. The Filled sites were located downstream of at least one VF. The Mined sites were located downstream of some mining activities but were not downstream of any VFs. The Filled/Residential sites were located downstream of at least one VF, and were also near residential areas. The Mined/Residential sites were located downstream of mining activity, and were also near residential areas. The additive sites were located on a mainstem of a watershed and were downstream of multiple VFs and VF-influenced streams.

2.3. Study Areas

2.3.1. Mud River Watershed

The headwaters of the Mud River are in Boone County, West Virginia, and flow northwest into Lincoln County, West Virginia. Although the headwaters of this watershed do not lie in the primary MTM/VF Region, there is a portion of the watershed that lies perpendicular to a five-mile strip of land in which mining activities are occurring. From the headwaters to the northwestern boundary of the primary MTM/VF Region, the watershed lies in the Cumberland Mountains of the Central Appalachian Plateau. The physiography is unglaciated, dissected hills and mountains with steep slopes and very narrow ridge tops and the geology is Pennsylvania sandstone, siltstone, shale, and coal of the Pottsville Group and Allegheny Formation (Woods et al. 1999). The primary land use is forest with extensive coal mining, logging, and gas wells. Some livestock farms and scattered towns exist in the wider valleys. Most of the low-density residential land use is concentrated in the narrow valleys (Green et al. 2000 Draft).

The U.S. EPA Region 3 sampled ten sites in the Mud River Watershed (Figure 2-2, Table 2-1). Brief descriptions of these sites are given below and more complete descriptions are given in Green et al. (2000 Draft). Site MT01 was established on the Mud River and the major disturbances at this site are a county road and residences. There also have been a few historical mining activities conducted upstream of site MT01. Site MT02 was established on Rush Patch Branch upstream of all residences and farms. While there is no history of mining in this sub-watershed, there is evidence of logging and gas well development. Site MT03 was established well above the mouth of Lukey Fork. Logging is the only known disturbance upstream of this site. Site MT13 was established on the Spring Branch of Ballard Fork. Other than historical logging activity, there is very little evidence of human disturbance associated with this site. Site MT14 was established on Ballard Fork. It is located downstream of eight VFs for which the mining permits were issued in 1985, 1988 and 1989. Site MT15 was established on Stanley Fork, located downstream of six VFs for which mining permits were issued in 1988, 1989, 1991, 1992 and 1995. Site MT24 was established in a sediment control structure on top of the mining operation located in the Stanley Fork sub-watershed. Site MT18 was established on Sugartree Branch. It was located downstream of two VFs for which the mining permits were

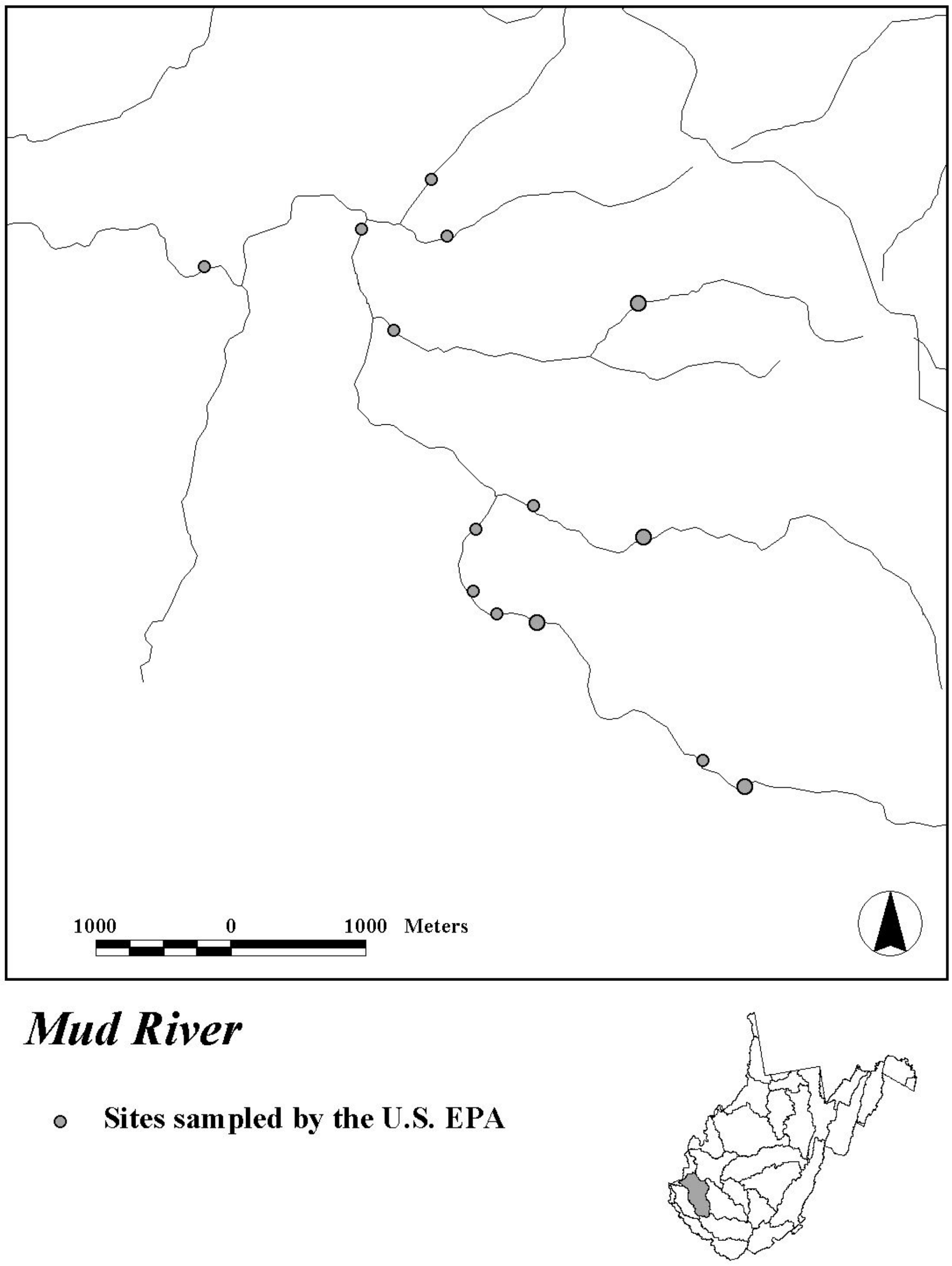


Figure 2-2. Sites sampled in the Mud River Watershed.

Table 2-1. Sites sampled in the Mud River Watershed.

Site ID/Organization	Stream Name	EIS Class
U.S. EPA Region 3		
MT01	Mud River	Mined/Residential
MT02	Rushpatch Branch	Unmined
MT03	Lukey Fork	Unmined
MT13	Spring Branch	Unmined
MT14	Ballard Fork	Filled
MT15	Stanley Fork	Filled
MT24	Unnamed Trib. to Stanley Fork	Sediment Control Structure
MT18	Sugartree Branch	Filled
MT23	Mud River	Filled/Residential
MT16	Unnamed Trib. to Sugartree Branch	Mined

issued in 1992 and 1995. Site MT23 was established on the Mud River downstream of mining activities. These activities include active and inactive surface mines and one active underground mine. In the spring of 2000, Site MT16 was established on an unnamed tributary to Sugartree Branch. This site was downstream of historical surface mining activities, but was not downstream of any VFs (Green et al. 2000 Draft).

2.3.2. Spruce Fork Watershed

The Spruce Fork Watershed drains portions of Boone and Logan Counties, West Virginia. The stream flows in a northerly direction to the town of Madison, West Virginia where it joins Pond Fork to form the Little Coal River. Approximately 85 to 90% of the watershed resides in the primary MTM region. Only the northwest corner of the watershed lies outside of this region. The entire watershed lies in the Cumberland Mountains sub-ecoregion (Woods et al. 1999). The watershed has been the location of surface and underground mining for many years, therefore, much of the watershed has been disturbed (Green et al. 2000 Draft).

The U.S. EPA Region 3 sampled eight sites in the Spruce Fork Watershed (Figure 2-3, Table 2-2). Brief descriptions of these sites are given below and more complete descriptions are given in Green et al. (2000 Draft). The U.S. EPA Region 3 Site MT39 was established on White Oak Branch and no mining activities existed in this area. Site MT40 was established on Spruce Fork. It is located downstream of seven known surface mining VFs and three VFs associated with refuse disposal. Site MT42 was established on Oldhouse Branch, located upstream of all residences and there is no known history of mining activities in this area. Site MT45 was

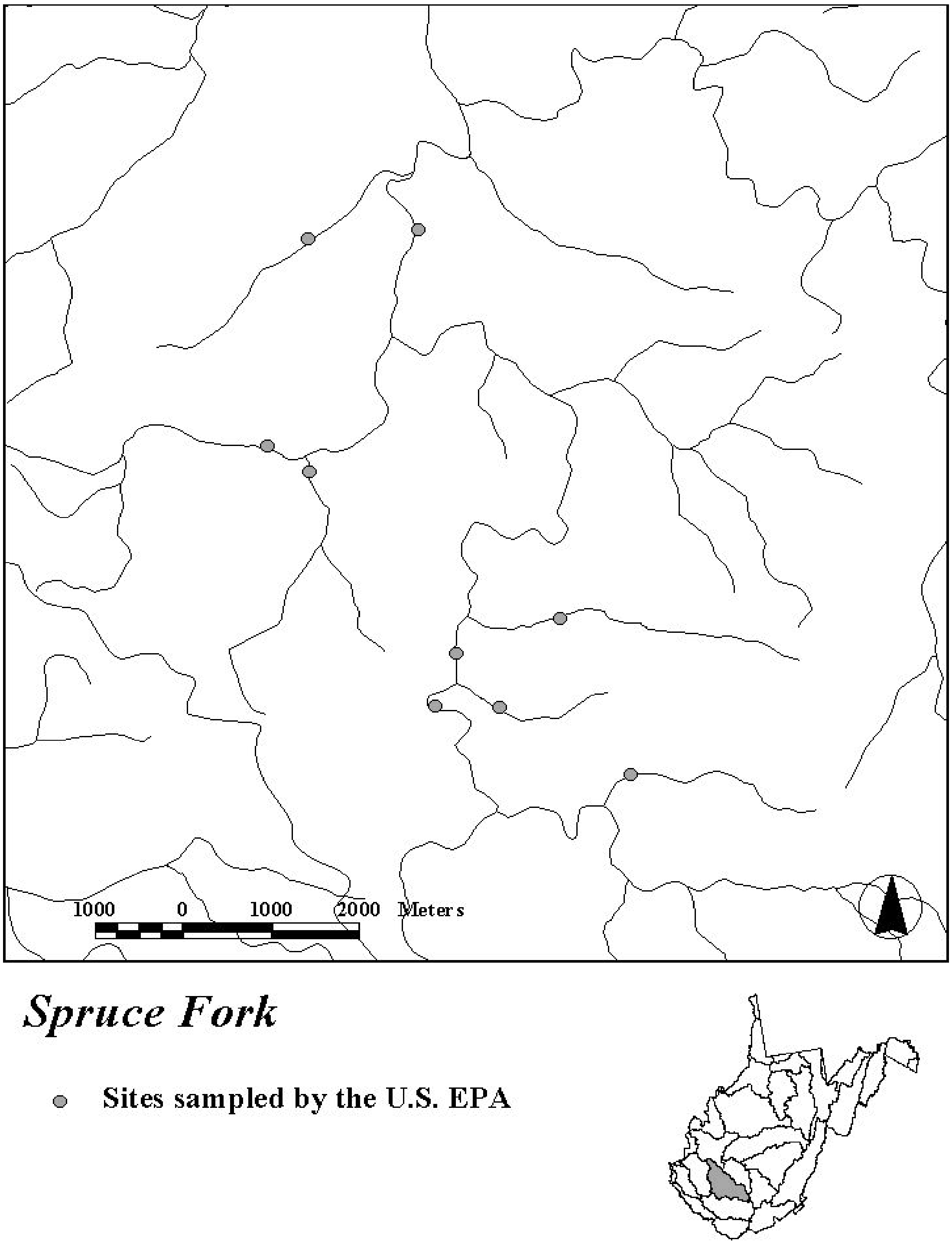


Figure 2-3. Sites sampled in the Spruce Fork Watershed.

Table 2-2. Sites sampled in the Spruce Fork Watershed.

Site ID/Organization	Stream Name	EIS Class
U.S. EPA Region 3		
MT39	White Oak Branch	Unmined
MT40	Spruce Fork	Filled/Residential
MT42	Oldhouse Branch	Unmined
MT45	Pigeonroost Branch	Mined
MT32	Beech Creek	Filled
MT34B	Left Fork	Filled
MT48	Spruce Fork	Filled/Residential
MT25B	Rockhouse Creek	Filled

established on Pigeonroost Branch. This site was located upstream of all residences but downstream of contour mining activities that occurred between 1987 and 1989. Site MT32 was established on Beech Creek. It was located downstream of five VFs and surface and underground mining activities. Site MT34B was established on the Left Fork of Beech Creek. It was located downstream of VFs and surface and underground mining activities. Site MT48 was established on Spruce Fork just upstream of Rockhouse Creek. There are known to be 22 VFs and several small communities upstream of this site. Site MT25B was established on Rockhouse Creek, located downstream of a sediment pond and a very large VF (Green et al. 2000 Draft).

2.3.3. Clear Fork Watershed

Clear Fork flows north toward its confluence with Marsh Fork where they form the Big Coal River near Whitesville, West Virginia. The entire watershed lies within Raleigh County, West Virginia within the Cumberland Mountains sub-ecoregion and, except for a very small portion, it lies within the primary MTM region (Woods et al. 1999). The coal mining industry has been active in this watershed for many years. Both surface and underground mining have occurred in the past and presently continue to be mined. There were no unmined sites sampled from this watershed (Green et al. 2000 Draft).

The U.S. EPA Region 3 sampled eight sites in the Clear Fork Watershed (Figure 2-4, Table 2-3). Brief descriptions of these sites are given below and more complete descriptions are given in Green et al. (2000 Draft). The U.S. EPA Region 3 Site MT79 was established on Davis Fork. It was located downstream of mining activities. Site MT78 was established on Raines Fork. It was located downstream of historical contour and underground mining. Site MT81 was

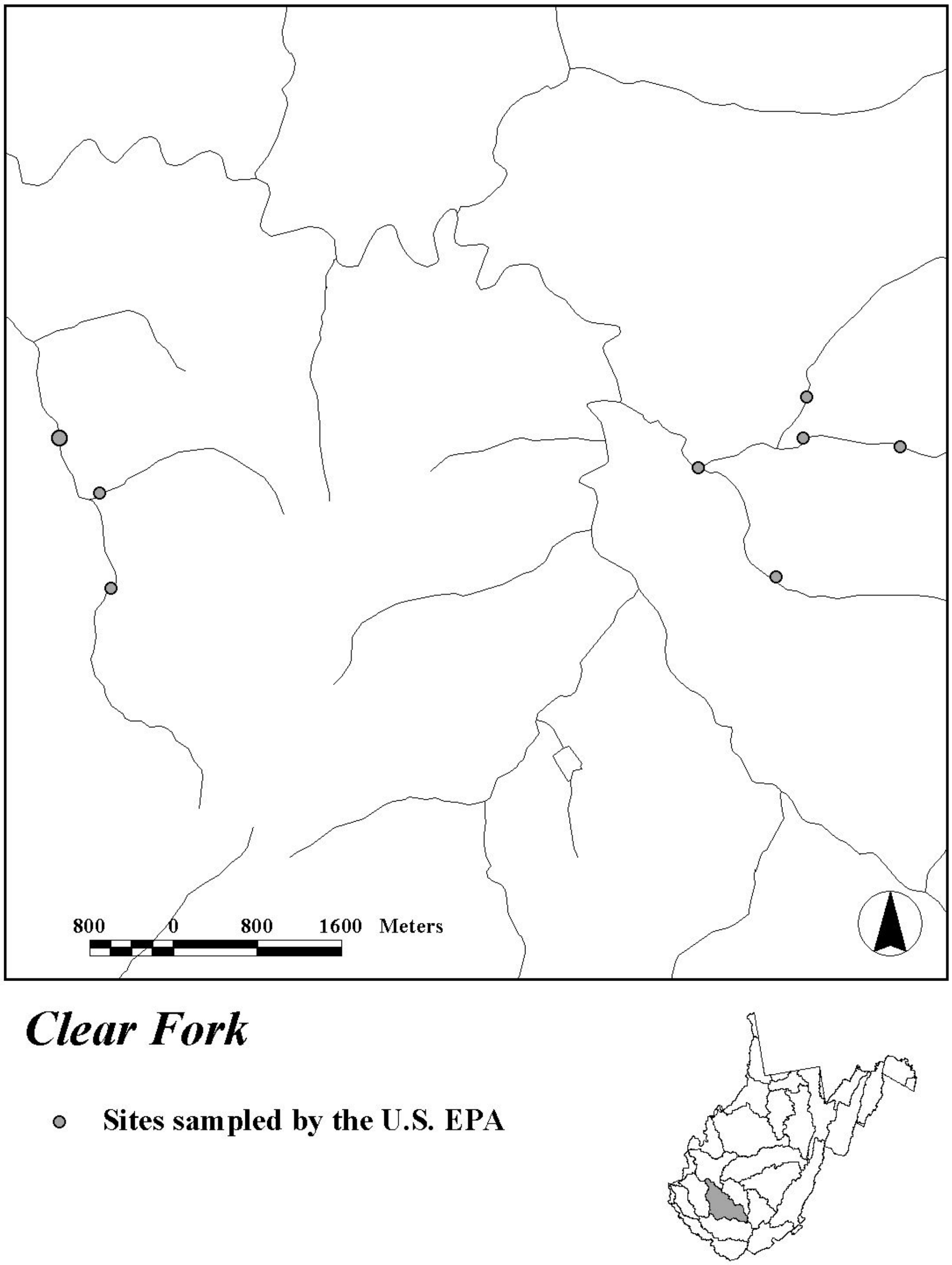


Figure 2-4. Sites sampled in the Clear Fork Watershed.

Table 2-3. Sites sampled in the Clear Fork Watershed.

Site ID/Organization	Stream Name	EIS Class
U.S. EPA Region 3		
MT79	Davis Fork	Mined
MT78	Raines Fork	Mined
MT81	Sycamore Creek	Mined
MT75	Toney Fork	Filled/Residential
MT70	Toney Fork	Filled/Residential
MT69	Ewing Fork	Mined/Residential
MT64	Buffalo Fork	Filled
MT62	Toney Fork	Filled/Residential

established on Sycamore Creek. It was located downstream of historical contour and underground mining and it is downstream of a plant that treats mine effluent. Site MT75 was established on Toney Fork. It was located downstream of five VFs, MTM activities and numerous residences. Site MT70 was established approximately 1 km (0.6 mi) downstream of Site MT75. It was located downstream of six VFs, MTM activities and numerous residences. This site was only sampled during autumn 1999 and winter and spring 2000. Site MT69 was established on Ewing Fork. It was located downstream of some historical contour and underground mining activities and a residence. Site MT64 was established on Buffalo Fork. It was located downstream of historical contour mining, current MTM activities, five VFs and a small amount of pasture. Site MT62 was established on Toney Fork. It was located downstream of 11 VFs, numerous residences and a small amount of pasture (Green et al. 2000 Draft).

2.3.4. Twentymile Creek Watershed

Twentymile Creek drains portions of Clay, Fayette, Kanawha, and Nicholas Counties, West Virginia. It generally flows to the southwest where it joins the Gauley River at Belva, West Virginia. Except for a small area on the western edge of the watershed, it is within the primary MTM region and the entire watershed lies within the Cumberland Mountains sub-ecoregion (Woods et al. 1999). Upstream of Vaughn, West Virginia, the watershed is uninhabited and logging, mining, and natural gas extracting are the primary activities. The majority of the mining activity has been conducted recently. Downstream of Vaughn, there are numerous residences and a few small communities (Green et al. 2000 Draft).

The U.S. EPA Region 3 sampled seven sites in the Twentymile Creek Watershed (Figure 2-5, Table 2-4). Brief descriptions of these sites are given below and more complete description

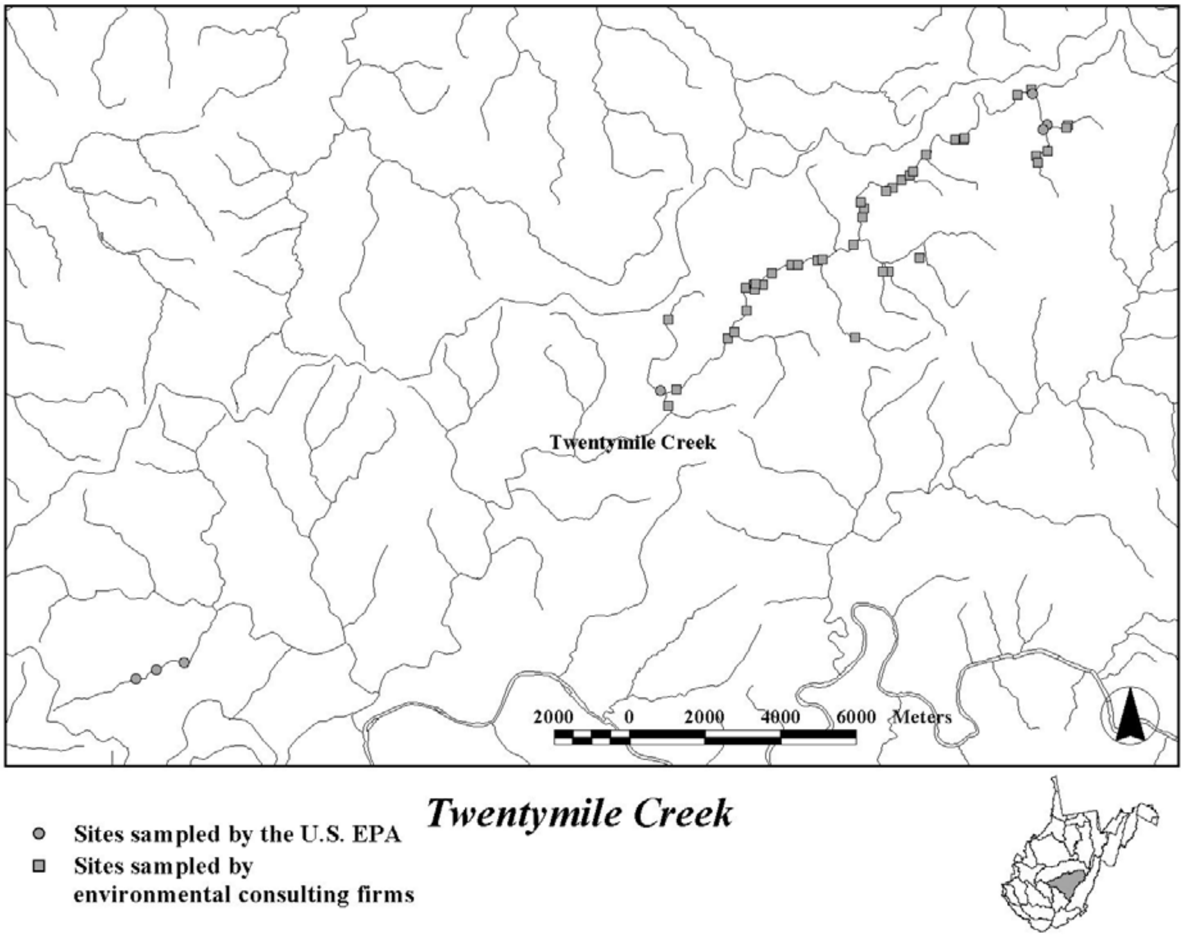


Figure 2-5. Sites sampled in the Twentymile Creek Watershed.

are given in Green et al. (2000 Draft). The U.S. EPA Region 3 Site MT95 was established on Neil Branch. There were no known disturbances upstream of this site. Site MT91 was established on Rader Fork. The only known disturbance to this site was a road with considerable coal truck traffic. Site MT87 was established on Neff Fork downstream of three VFs and a mine drainage treatment plant. Site MT86 was located on Rader Fork downstream of Site MT91 and Neff Fork and it was, therefore, downstream of three VFs and a mine drainage treatment plant. Site MT103 was established on Hughes Fork. It was downstream of six VFs. Site MT98 was established on Hughes Fork. It was downstream of Site MT103 and eight VFs. Site MT104 was established on Hughes Fork. It was downstream of Site MT103, Site MT98, eight VFs and a sediment pond (Green et al. 2000 Draft).

Table 2-4. Sites sampled in the Twentymile Creek Watershed. Equivalent sites are noted parenthetically.

Site ID/Organization	Stream Name	EIS Class
U.S. EPA Region 3		
MT95 (=Neil-5)	Neil Branch	Unmined
MT91	Rader Fork	Unmined
MT87 (=Rader-4)	Neff Fork	Filled
MT86 (=Rader-7)	Rader Fork	Filled
MT103	Hughes Fork	Filled
MT98	Hughes Fork	Filled
MT104	Hughes Fork	Filled
BMI		
Rader 8	Twentymile Creek	Additive
Rader 9	Twentymile Creek	Additive
PMC-TMC-36	Twentymile Creek	Additive
PMC-TMC-35	Twentymile Creek	Additive
PMC-TMC-34	Twentymile Creek	Additive
PMC-TMC-33	Twentymile Creek	Additive
PMC-TMC-31	Twentymile Creek	Additive
PMC-TMC-30	Twentymile Creek	Additive
PMC-TMC-29	Twentymile Creek	Additive
PMC-TMC-28	Twentymile Creek	Additive
PMC-TMC-27	Twentymile Creek	Additive
PMC-TMC-26	Twentymile Creek	Additive
PMC-7	Twentymile Creek	Additive
PMC-6	Twentymile Creek	Additive
PMC-5	Twentymile Creek	Additive
PMC-TMC-4	Twentymile Creek	Additive
PMC-TMC-5	Twentymile Creek	Additive
PMC-TMC-314	Twentymile Creek	Additive
PMC-TMC-2	Twentymile Creek	Additive
PMC-TMC-1	Twentymile Creek	Additive

Continued

Table 2-4. Continued.

Site ID/Organization	Stream Name	EIS Class
BMI (Continued)		
PMC-HWB-1	Twentymile Creek	Additive
PMC-HWB-2	Twentymile Creek	Additive
Neil-6 (=Fola 48)	Twentymile Creek	Additive
Neil-7 (=Fola 49)	Twentymile Creek	Additive
Neil-2 (=Fola 53)	Neil Branch	Unmined
Neil-5 (=MT95)	Neil Branch	Unmined
Rader-1	Laurel Run	Unmined
Rader-2	Rader Fork	Unmined
Rader-3	Trib. to Rader	Unmined
Rader-4 (=MT87)	Neff Fork	Filled (2)
Rader-5	Neff Fork	Filled (2)
Rader-6	Trib. to Neff	Filled (1)
Rader-7 (=MT86)	Rader Fork	Filled (2)
PMC-1	Sugarcamp Branch	Filled (1)
PMC-11	Right Fork	Filled (1)
PMC-12	Road Fork	Filled (1)
PMC-15	Tributary to Robinson Fork.	Filled (1)
POTESTA		
Fola 33	Twentymile Creek	Additive
Fola 36	Twentymile Creek	Additive
Fola 37	Twentymile Creek	Additive
Fola 38	Twentymile Creek	Additive
Fola 48 (=Neil-6)	Twentymile Creek	Additive
Fola 49 (=Neil-7)	Twentymile Creek	Additive
Fola 39	Peachorchard Branch	Filled (2 small)
Fola 40	Peachorchard Branch	Filled (1 small)
Fola 45	Peachorchard Branch	Unmined
Fola 53 (=Neil-2)	Neil Branch	Unmined

2.3.5. Island Creek Watershed

Island Creek generally flows north toward Logan, West Virginia where it enters the Guyandotte River. The entire watershed is confined to Logan County. With the exception of the northern portion, the watershed lies within the primary MTM region and the entire watershed lies within the Cumberland Mountains sub-ecoregion (Woods et al. 1999). Extensive underground mining has occurred in the watershed for many years. As the underground reserves have been depleted and the economics of the area have changed, surface mining has played a larger role in the watershed (Green et al. 2000 Draft).

The U.S. EPA Region 3 sampled eight sites in the Island Creek Watershed (Figure 2-6, Table 2-5). Brief descriptions of these sites are given below and more complete descriptions are given in Green et al. (2000 Draft). The U.S. EPA Region 3 Site MT50 was located on Cabin Branch in the headwaters of the sub-watershed and upstream of any disturbances. Site MT51 was also established on Cabin Branch located downstream of Site MT50 and a gas well. Site MT107 was established on Left Fork in the spring of 2000, located upstream of the influence of VFs. Site MT52 was established near the headwaters of Cow Creek. It was located upstream of VFs, but downstream of an underground mine entrance, a small VF and a sediment pond. Site MT57B was established on Hall Fork for sampling in the spring and summer 1999. It was located downstream of a sediment pond and a VF. In the autumn 1999, Site MT57 was established near the mouth of Hall fork. It was farther downstream than Site MT57B and was downstream of a sediment pond and a VF. Site MT60 was established on Left Fork, downstream of Site MT107. It was located downstream of two existing VFs and three proposed VFs. Site MT55 was established on Cow Creek, downstream of Site MT52. It was located downstream of four VFs associated with MTM, one VF associated with underground mining, residences, a log mill, orchards, vineyards, cattle, and a municipal sewage sludge disposal site (Green et al. 2000 Draft).

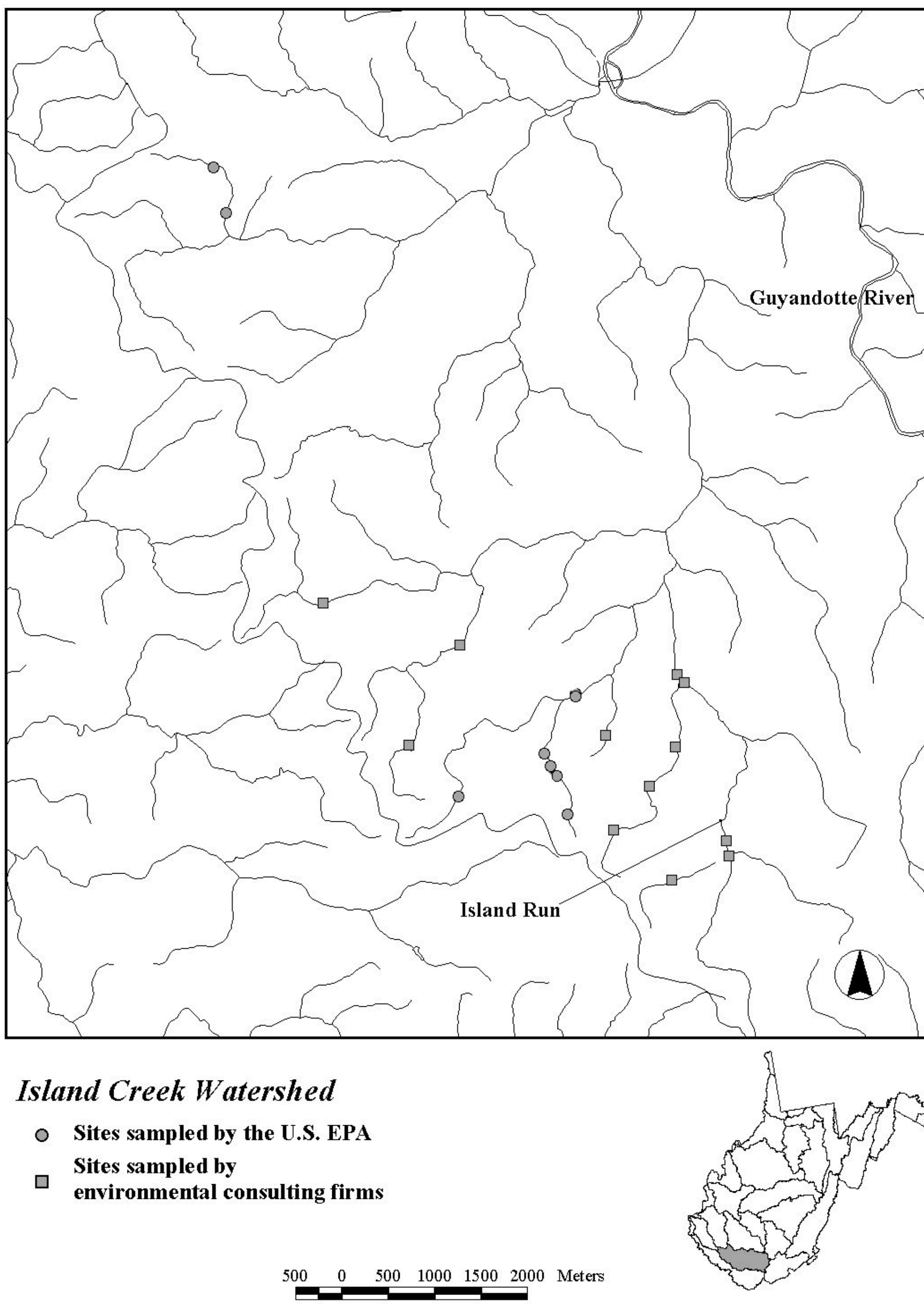


Figure 2-6. Sites sampled in the Island Creek Watershed.

Table 2-5. Sites sampled in the Island Creek Watershed.

Site	Stream Name	EIS Class
U.S. EPA Region 3		
MT50	Cabin Branch	Unmined
MT51	Cabin Branch	Unmined
MT107	Left Fork	Unmined
MT52	Cow Creek	Filled
MT57B	Hall Fork	Filled
MT57	Hall Fork	Filled
MT60	Left Fork	Filled
MT55	Cow Creek	Filled/Residential
BMI		
Mingo 34		Filled (1)
Mingo 41		Filled (2)
Mingo 39		Filled (1) + old mining
Mingo 16		Unmined
Mingo 11		Unmined
Mingo 2		Unmined
Mingo 86		Unmined
Mingo 62		Unmined
Mingo 38	Island Creek	Additive
Mingo 24	Island Creek	Additive
Mingo 23	Island Creek	Additive

2.3.6. Twelvepole Creek Watershed

The East Fork of the Twelvepole Creek Watershed drains portions of Mingo, Lincoln, and Wayne Counties, West Virginia. The stream flows northwest to the town of Wayne, West Virginia where it joins the West Fork of Twelvepole Creek then continues to flow on into the Ohio River at Huntington, West Virginia. The East Fork of Twelvepole Creek is impounded by East Lynn Lake near Kiahsville, West Virginia in Wayne County (West Virginia DEP, Personal Communication).

The East Fork of the Twelvepole Creek Watershed encompasses approximately 445 km^2 (172 mi^2) of drainage area and is 93.3% forested. Prior to 1977, very little mining had occurred in the watershed south of East Lynn Lake. Since 1987, several surface mining operations have been employed in the Kiah Creek and the East Fork of Twelvepole Creek watersheds (Critchley 2001). Currently, there are 23 underground mining, haul road and refuse site permits, and 21 surface mining permits in the watershed (West Virginia DEP, Personal Communication).

REIC has conducted biological evaluations in the East Fork of the Twelvepole Creek Watershed since 1995. Five stations have been sampled on Kiah Creek (Figure 2-7, Table 2-6). Station BM-003A was located in the headwaters of Kiah Creek, upstream from surface mining and residential disturbances. Station BM-003 was located near the border of Lincoln and Wayne Counties and it was downstream from several surface mining operations and several residential disturbances. Station BM-004 was located on Kiah Creek downstream from the surface mining operations on Queens Fork and Vance Branch, near the confluence of Jones Branch, downstream from Trough Fork, and downstream of residential disturbances. Station BM-004A was located downstream from the confluence of Big Laurel Creek, surface mining operations and residential disturbances.

Two stations were sampled in Big Laurel Creek (Figure 2-7, Table 2-6). This tributary has only residential disturbances in its watershed. Station BM-UBLC was located near the headwaters of Big Laurel Creek. Station BM-DBLC was located near the confluence of Big Laurel Creek with Kiah Creek.

Eight stations were sampled on the East Fork of Twelvepole Creek (Figure 2-7, Table 2-6). Station BM-001A was located just downstream from confluence of McCloud Branch and was downstream of a residential disturbance. Station BM-001C was located downstream of the confluence of Laurel Branch which currently has a VF, additional proposed VFs, and residences. Station BM-001B was located downstream of the confluence of Wiley Branch which has residences, numerous current VFs and additional VFs under construction or being proposed. Station BM-001 was located upstream from the confluence of Bluewater Branch but downstream from the Wiley Branch and Laurel Branch surface mining operations and residences. Station BM-010 was downstream from the Franks Branch mining operation and residences. Station BM-011 was located downstream from the Maynard Branch operations and residences. Station BM-002 was located downstream from the Devil Trace surface mining operation and residences. Station BM-002A was located downstream of Milam Creek and all mining operations and residences in this sub-watershed.

Two stations were located in Milam Creek, a tributary of the East Fork of Twelvepole Creek (Figure 2-7, Table 2-6). Milam Creek has no mining operations or residential disturbances in its watershed. Station BM-UMC was located near the headwaters of Milam Creek and station BM-DMC was located near the confluence of Milam Creek with the East Fork of Twelvepole Creek.

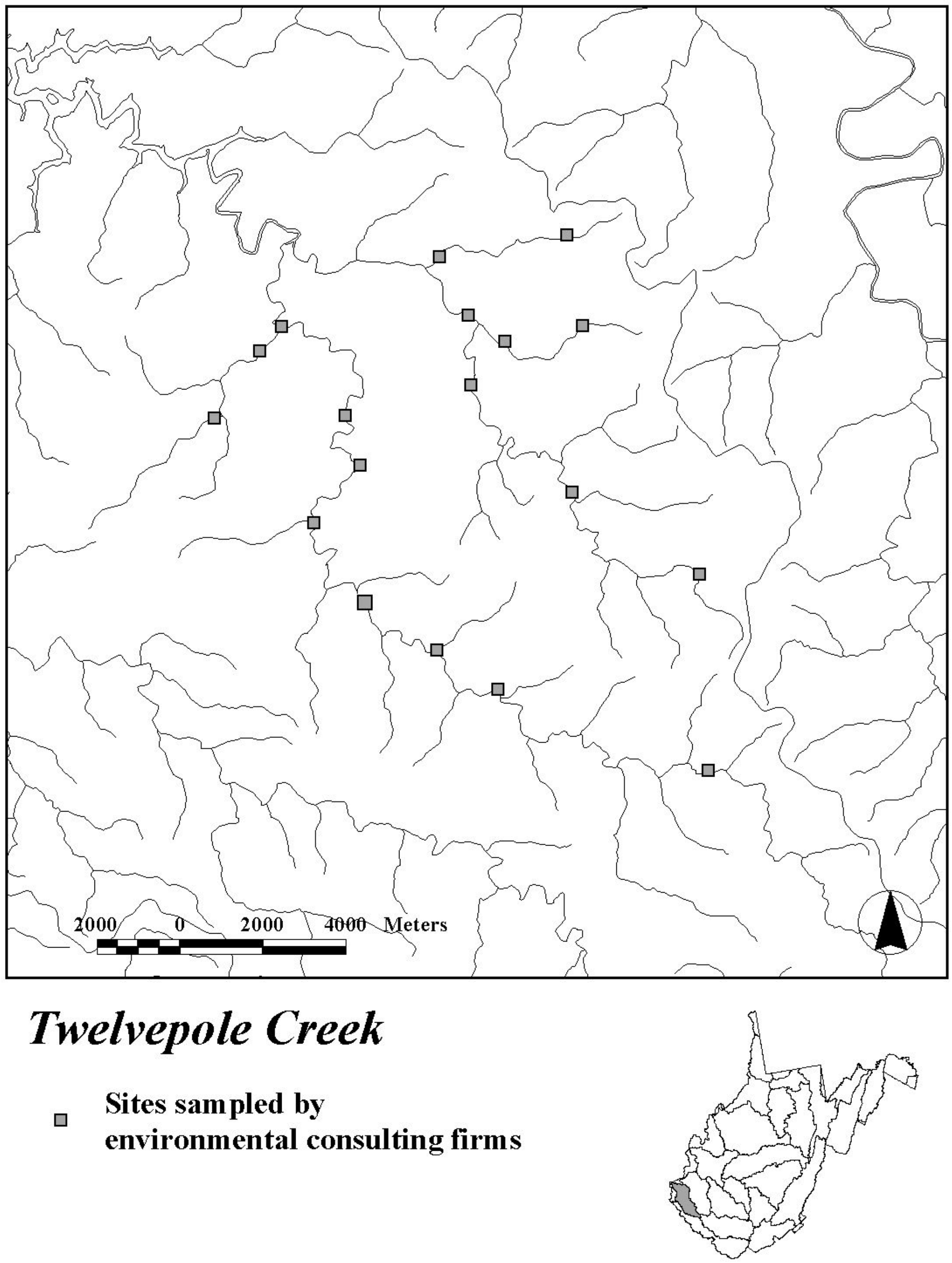


Figure 2-7. Sites sampled in the Twelvepole Creek Watershed.

Table 2-6. Sites sampled in the Twelvepole Creek Watershed. Equivalent sites are noted parenthetically.

Site ID/Organization	Stream Name	EIS Class
REIC		
BM-003A	Kiah Creek	Additive
BM-003	Kiah Creek	Additive
BM-004	Kiah Creek	Additive
BM-004A	Kiah Creek	Additive
BM-DBLC	Big Laurel Creek	Unmined
BM-UBLC	Big Laurel Creek	Unmined
BM-001A	Twelvepole Creek	Additive
BM-001C	Twelvepole Creek	Additive
BM-001B	Twelvepole Creek	Additive
BM-001	Twelvepole Creek	Additive
BM-010	Twelvepole Creek	Additive
BM-011	Twelvepole Creek	Additive
BM-002	Twelvepole Creek	Additive
BM-002A	Twelvepole Creek	Additive
BM-UMC	Milam Creek	Unmined
BM-DMC	Milam Creek	Unmined
BM-005	Trough Fork	Additive
BM-006	Trough Fork	Additive

2.4. Data Collection Methods

The data for this study were generated by five different organizations (i.e., U.S. EPA Region 3, PSU, BMI, POTESTA and REIC). The methods used to collect each of the four different types of data (i.e., habitat, water quality, fish assemblage and macroinvertebrate assemblage) are described below. This information is summarized in tabular form in Appendix A.

2.4.1. Habitat Assessment Methods

2.4.1.1. U.S. EPA Region 3 Habitat Assessment

The U.S. EPA Region 3 used the RBP (Barbour et al. 1999) to collect habitat data at each site. Although some parameters require observations of a broader section of the catchment area, the habitat data were primarily collected in a 100-m reach that includes the portion of the stream where biological data (i.e., fish and macroinvertebrate samples) were collected. The RBP habitat assessment evaluates ten parameters (Appendix A).

The U.S. EPA Region 3 measured substrate size and composition in order to help determine if excessive sediment was causing any biological impairments (Kaufmann and Robison 1998). Numeric scores were assigned to the substrate classes that are proportional to the logarithm of the midpoint diameter of each size class (Appendix A).

2.4.1.2. BMI Habitat Assessment

The Standard Operating Procedures (SOPs) submitted by BMI make no mention of habitat assessment methods.

2.4.1.3. POTESTA Habitat Assessment

POTESTA collected physical habitat data using methods outlined in Kaufmann et al. (1999) or in Barbour et al. (1999, Appendix A). The habitat assessments were performed on the same reaches from which biological sampling was conducted. A single habitat assessment form was completed for each sampling site. This assessment form incorporated features of the selected sampling reach as well as selected features outside the reach but within the catchment area. Habitat evaluations were first made on in-stream habitat, followed by channel morphology, bank structural features, and riparian vegetation.

2.4.1.4. REIC Habitat Assessment

The SOPs submitted by REIC make no mention of habitat assessment methods.

2.4.2. Water Quality Assessment Methods

2.4.2.1. U.S. EPA Water Quality Assessment

The U.S. EPA Region 3 measured conductivity, pH, temperature and dissolved oxygen (DO) *in situ* and the flow rate of the stream at the time of sampling. Each of these measurements was made once at each site during each field visit. The U.S. EPA Region 3 also collected water samples for laboratory analyses. These samples were analyzed for the parameters given in Table 2-7.

2.4.2.2. BMI Water Quality Assessment

The SOPs submitted by BMI make no mention of water quality assessment methods.

2.4.2.3. POTESTA Water Quality Assessment

POTESTA measured conductivity, pH, temperature and DO *in situ.* These measurements were taken once upstream from each biological sampling site, and were made following the protocols outlined in U.S. EPA (1979). The stream flow rate was also measured at or near each sampling point. One of the three procedures (i.e., velocity-area, time filling, or neutrally buoyant object) outlined in Kaufmann (1998) was used at each site. POTESTA also collected water samples at each site directly upstream of the location of the biological sampling. These samples were analyzed in the laboratory for the suite of analytes listed in Table 2-7.

2.4.2.4. REIC Water Quality Assessment

REIC recorded water body characteristics (i.e., size, depth and flow) and site location at each site. Grab samples were collected and delivered to the laboratory for analysis. The SOPs submitted by REIC make no mention of which analytes were measured in the laboratory.

2.4.3. Fish Assemblage Methods

2.4.3.1. PSU Fish Assemblage Assessment

The PSU, in consultation with personnel from U.S. EPA Region 3, sampled fish assemblages at 58 sites in West Virginia. The fish sampling procedures generally followed those in McCormick and Hughes (1998). Fish were collected by making three passes using a backpack electrofishing unit. Each pass proceeded from the downstream end of the reach to the upstream

Table 2-7. Parameters used by each organization for lab analyzed water samples.

Parameter	Organizations			
	U.S. EPA	BMI	POTESTA	REIC
Acidity	Yes	Unknown	Yes	Unknown
Alkalinity	Yes	Unknown	Yes	Unknown
Chloride	Yes	Unknown	Yes	Unknown
Hardness	Yes	Unknown	Yes	Unknown
Nitrate(NO_3) + Nitrite (NO_2)	Yes	Unknown	Yes	Unknown
Sulfate	Yes	Unknown	Yes	Unknown
Total Suspended Solids (TSS)	Yes	Unknown	Yes	Unknown
Total Dissolved Solids (TDS)	Yes	Unknown	Yes	Unknown
Total Organic Carbon (TOC)	Yes	Unknown	Yes	Unknown
Coarse Particulate Organic Matter (CPOM)	No	Unknown	Yes	Unknown
Fine Particulate Organic Matter (FPOM)	No	Unknown	Yes	Unknown
Total Dissolved Organic Carbon (TDOC)	Yes	Unknown	No	Unknown
Total Aluminum	Yes	Unknown	Yes	Unknown
Dissolved Aluminum	Yes	Unknown	Yes	Unknown
Total Antimony	Yes	Unknown	Yes	Unknown
Total Arsenic	Yes	Unknown	Yes	Unknown
Total Barium	Yes	Unknown	No	Unknown
Total Beryllium	Yes	Unknown	Yes	Unknown
Total Cadmium	Yes	Unknown	Yes	Unknown
Total Calcium	Yes	Unknown	Yes	Unknown
Total Chromium	Yes	Unknown	Yes	Unknown
Total Cobalt	Yes	Unknown	No	Unknown
Total Copper	Yes	Unknown	Yes	Unknown
Total Iron	Yes	Unknown	Yes	Unknown

(Continued)

Table 2-7. Continued.

Parameter	Organizations			
	U.S. EPA	BMI	POTESTA	REIC
Dissolved Iron	Yes	Unknown	Yes	Unknown
Total Lead	Yes	Unknown	Yes	Unknown
Total Magnesium	Yes	Unknown	Yes	Unknown
Total Manganese	Yes	Unknown	Yes	Unknown
Dissolved Manganese	Yes	Unknown	Yes	Unknown
Total Mercury	Yes	Unknown	Yes	Unknown
Total Nickel	Yes	Unknown	Yes	Unknown
Total Potassium	Yes	Unknown	Yes	Unknown
Total Phosphorous	Yes	Unknown	Yes	Unknown
Total Selenium	Yes	Unknown	Yes	Unknown
Total Silver	Yes	Unknown	Yes	Unknown
Total Sodium	Yes	Unknown	Yes	Unknown
Total Thallium	Yes	Unknown	Yes	Unknown
Total Vanadium	Yes	Unknown	No	Unknown
Total Zinc	Yes	Unknown	Yes	Unknown

end of the reach. Block nets were used only when natural barriers (i.e., shallow riffles) were not present. The fish collected from each pass were kept separate. Fish were identified to the species level and enumerated. The standard length of each fish was measured to the nearest mm and each fish was weighed to the nearest 0.01 g.

2.4.3.2. BMI Fish Assemblage Assessment

The SOPs submitted by BMI make no mention of fish assemblage assessment methods.

2.4.3.3. POTESTA Fish Assemblage Assessment

POTESTA collected fish by using the three-pass depletion method of Van Deventer and Platts (1983) with a backpack electrofishing unit. Each of the three passes proceeded from the downstream end of the reach to the upstream end of the reach. The fish collected from each pass were kept separate. Additional passes were made if the numbers of fish did not decline during the two subsequent passes. Game fish and rare, threatened or candidate (RTC) fish species were identified, their total lengths were recorded to the nearest mm, and their weights were recorded to the nearest g. With the exception of small game and non-RTC fish, the captured fish were released. Small game fish and non-RTC fish that were collected during each pass were preserved separately and transported to the laboratory for analysis. Preserved fish were identified and weighed to the nearest g.

2.4.3.4. REIC Fish Assemblage Assessment Methods

REIC collected fish by setting block nets across the stream and perpendicular to the stream banks, then progressing upstream with a backpack electrofishing unit. The entire reach was surveyed three times. After each survey, all large fish were identified using guidelines given by Trautman (1981) and Stauffer et al. (1995). The total lengths of the fish were measured to the nearest mm and they were weighed to the nearest g. After all three passes were completed, the large fish were returned to the stream. Small fish which required microscopic verification of their identification were preserved and transported to the laboratory. Once in the laboratory, small fish were identified using guidelines given by Trautman (1981) and Stauffer et al. (1995). After identification, the total lengths of the fish were measured to the nearest mm, they were weighed to the nearest 0.1 g and their identifications were reconfirmed.

2.4.4. Macroinvertebrate Assemblage Methods

2.4.4.1. U.S. EPA Region 3 Macroinvertebrate Assemblage Assessment

The U.S. EPA's Region 3 used RBPs to assess benthic macroinvertebrate assemblages (Barbour et al. 1999). Samples were collected from riffles only. A 0.5 m wide rectangular dip net with 595-µm mesh was used to collect organisms in a 0.25 m^2 area upstream of the net. At each site, four samples were taken, and composited into a single sample, representing a total area sampled of approximately 1.0 m^2. The RBPs recommend the total area sampled to be 2.0 m^2 but that was reduced to 1.0 m^2 for this study due to the small size of the streams. Benthic macroinvertebrate samples were collected in each season except when there was not enough flow for sampling. Approximately 25% of the sites were sampled in replicate to provide information on within-season and within-site variability. These replicate samples were collected at the same time, usually from adjacent locations in the same riffle.

The samples collected by the U.S. EPA Region 3 were sub-sampled in the laboratory so that ⅛ of the composite samples were picked. All organisms in the sub-sample were identified to the family level, except for oligochetes and leeches, which were identified to the class level. Organisms were identified using published taxonomic references (i.e., Pennak 1989, Pecharsky et al. 1990, Stewart and Stark 1993, Merritt and Cummins 1996, Westfall and May 1996, Wiggins 1998).

2.4.4.2. BMI Macroinvertebrate Assemblage Methods

BMI collected samples using a kick net with a 0.5 m width and a 600 μm mesh size. The net was held downstream of the 0.25 m^2 area that was to be sampled. All rocks and debris that were in the 0.25 m^2 area were scrubbed and rinsed into the net and removed from the sampling area. Then, the substrate in the 0.25 m^2 area was vigorously disturbed for 20 seconds. This process was repeated four times at each sampling site and the four samples were composited into a single sample.

BMI also collected samples using a 0.09 m^2 (1.0 ft^2) Surber sampler with a 600 μm mesh size. The frame of the sampler was placed on the stream bottom in the area that was to be sampled. All large rocks and debris that were in the 1.0-ft^2 frame were scrubbed and rinsed into the net and removed from the sampling area. Then, the substrate in the 1.0 ft^2 frame was vigorously disturbed for 20 seconds. In autumn 1999 and spring 2000, no samples were collected with Surber samplers. In autumn 2000, six Surber samples were collected at each site, and in spring 2001, four Surber samples were collected. All Surber samples were kept separate.

In the laboratory, the samples were rinsed using a sieve with 700 μm mesh. All macroinvertebrates in the samples were picked from the debris. Each organism was identified to the taxa level specified in the project study plan.

2.4.4.3. POTESTA Macroinvertebrate Assemblage Assessment

POTESTA collected samples of macroinvertebrates using a composite of four 600 μm mesh kick net samples and following the U.S. EPA's RBPs (Barbour et al. 1999). For each of the four kick net samples, all large debris within a 0.25 m^2 area upstream of the kick net were brushed into the net. Then, the substrate in the 0.25 m^2 area was disturbed for 20 seconds. Once all four kick net samples were collected, they were composited into a single labeled jar.

POTESTA used Surber samplers to collect macroinvertebrate samples at selected sites. Surber samples were always collected in conjunction with kick net samples. At sites selected for quantitative sampling, a Surber sampler was placed on the stream bottom in a manner so that all sides were flat against the stream bed. Large cobble and gravel within the frame were thoroughly brushed and the substrate within the frame was disturbed for a depth of up to 7.6 cm

(3.0 in) with the handle of the brush. The sample was then placed in a labeled jar. The SOPs submitted by POTESTA make no mention of the area sampled or the number of samples collected with the Surber samplers.

In the laboratory, all organisms in the samples were identified by qualified freshwater macroinvertebrate taxonomists to the lowest practical taxonomic levels using Wiggins (1977), Stewart and Stark (1988), Pennak (1989) and Merritt and Cummins (1996). To ensure the quality of the identifications, 10% of all samples were re-picked and random identifications were reviewed.

2.4.4.4. REIC Macroinvertebrate Assemblage Assessment

REIC collected macroinvertebrate samples using a 600 µm mesh D-frame kick net. The kick net was positioned in the stream with the net outstretched with the cod end on the downstream side. The person using the net then used a brush to scrub any rocks within a 0.25 m^2 area in front of the net, sweeping dislodged material into the net. The person then either kicked up the substrate in the 0.25 m^2 area in front of the net or knelt and scrubbed the substrate in that area with one hand. The substrate was scrubbed or kicked for up to three minutes, with the discharged material being swept into the net. This procedure was repeated four times so that the total area sampled was approximately 1.0 m^2. Once collected, the four samples were composited into a single sample.

REIC also collected macroinvertebrate samples using Surber samplers with sampling areas of 0.09 m^2 (1 ft^2). These samplers were only used in areas where the water depth was less than 0.03 m (1 ft). The SOPs submitted by REIC make no mention of the mesh size used in the Surber samplers. The Surber sampler was placed in the stream, with the cod end of the net facing downstream. The substrate within the 1 ft^2 area was scrubbed for a period of up to three minutes and to a depth of approximately 7.62 cm (3 in). While being scrubbed, the dislodged material was swept into the net. After scrubbing was complete, rocks in the sampling area were checked for clinging macroinvertebrates. Once they had been removed, the material in the net was rinsed and the sample was deposited into a labeled sampling jar. Three Surber samples were collected at each site where they were used. These samples were not composited.

In the laboratory, REIC processed all samples individually. Samples were poured through a 250 µm sieve and rinsed with tap water. The sample was then split into quarters by placing it on a sub-sampling tray fitted with a 500 µm screen and spread evenly over the tray. The sample in the first quarter of the tray was removed, placed into petri dishes, and placed under a microscope so that all macroinvertebrates could be separated from the detritus. If too few organisms (this number is not specified in the SOPs submitted by REIC) were in the first quarter, then additional quarters were picked until enough organisms had been retrieved from the sample.

REIC used three experienced aquatic taxonomists to identify macroinvertebrates. They identified the organisms under microscopes to their lowest practical taxonomic level, usually Genus. Chironomids were often identified to the Family level and annelids were identified to the Class level. As taxonomic guides, REIC used Pennak (1989), Stewart and Stark (1993), Wiggins (1995), Merritt and Cummins (1996) and Westfall and May (1996).

3. DATA ANALYSES

3.1. Database Organization

3.1.1. Data Standardization

All of the methods used to collect and process fish samples were compatible, thus it was not necessary to standardize the fish data prior to analysis. However, there were differences among the methods used to collect and process the benthic macroinvertebrate data which made it necessary to standardize the macroinvertebrate data to eliminate potential biases before data analysis.

The benthic macroinvertebrate database was organized by sampling device (i.e., D-frame kick net or Surber sampler). Since not all organizations used Surber samplers and not all organizations that used Surber samplers employed the same methods (Section 2.4.4), Surber data were not used for the analyses in this report. All of the sampling organizations did use D-frame kick nets with comparable field methods to collect macroinvertebrate samples. Use of the data collected by D-frame kick net provides unbiased data with respect to the types, densities and relative abundances of organisms collected. However, while identifying organisms in the laboratory, the U.S. EPA sub-sampled 1/8 of the total material (with some exceptions noted in the data), REIC sub-sampled 1/4 of the total material (with some exceptions), and BMI and POTESTA counted the entire sample. To eliminate bias of the reported taxa richness data introduced by different sizes of sub-samples, all organism counts were standardized to a 1/8 sub-sample of the total original material. (Appendices A and E)

3.1.2. Database Description

3.1.2.1. Description of Fish Database

The fish database included 126 sampling events where the collection of a fish sample had been attempted and the location and watershed area were known. Of these, five were regional reference samples from Big Ugly Creek, outside of the study watersheds. Catchments with areas of less than 2.0 km^2 and samples with fewer than ten fish were excluded from the analysis (section 4.1.1). A summary of the remaining 99 samples is shown in Table 3-1.

The Mined/Residential EIS Class consisted of only two samples. Due to insufficient sample size for adequate statistical analysis, this class was eliminated.

Table 3-1. Number of fish sites and samples in the study area, by EIS class and watershed. The first numbers in the cells represent the number of sites and the numbers in parentheses represent the numbers of samples.

Watershed	Unmined	Filled	Mined	Filled/Res	Additive	Total
Mud River	3, (4)	4, (8)		1, (3)	1, (2)	9, (17)
Island Creek	1, (1)	2, (3)		2, (2)	2, (2)	7, (8)
Spruce Fork	1, (1)	3, (3)	1, (1)	3, (3)	1, (1)	9, (9)
Clear Fork		1, (1)	3, (3)	3, (3)		7, (7)
Twenty Mile Creek	5, (5)	7, (7)			7, (16)	19, (28)
Twelvepole Creek[1]	4, (6)				12, (24)	16, (30)
Total	14, (17)	17, (22)	4, (4)	9, (11)	23, (45)	67, (99)

[1]All sites in Twelvepole Creek were sampled by REIC; and were Additive and Unmined only.

3.1.2.2. Description of Macroinvertebrate Database

A total of 282 macroinvertebrate samples were collected from 66 sites in six watersheds (Table 3-2). The samples from sites in the Mined/Residential EIS class were removed from the analysis because there were too few sites (i.e., $n < 3$) to conduct statistical comparisons.

The U.S. EPA Region 3 collected a duplicate sample from the same site, on the same day, 42 different times, in five of the six sampled watersheds (i.e., no duplicate samples were taken from the Twelvepole Creek Watershed). The WVSCI, the total # of families, and the total number of EPT were highly correlated for duplicate samples (Table 3-3). Green et al. (2000) found similar results with raw metric scores. Because of these correlations and in order to avoid inflating the sample size, the only U.S. EPA Region 3 duplicate samples used for analyses were those that were labeled Replicate Number 1.

One site in Twentymile Creek was sampled by more than one organization the same season (i.e., autumn 2000 and winter 2001). To avoid sample size inflation, the means of the sample values were used for each season, thereby reducing the total number of samples. The means were used instead of the values from one of the samples because the samples were collected between three and five weeks apart. The U.S. EPA and two other organizations sampled the same site in the autumn 1999 and the winter 2000. In this case, the U.S. EPA data were used because these data did not require making a correction for sub-sampling.

Table 3-2. Number of sites and D-frame kick net samples available in each watershed and

in each EIS class.

Watershed	Unmined		Filled		Filled/ Residential		Mined		Mined/ Residential[1]		Total	
	Site	Samp	Site	Samp	Site	Samp	Site	Samp	Site	Samp	Site	Samp
Mud River	3	11	3	19	1	6	1	1	1	5	9	42
Island Creek	7	13	6	21	1	6	1	1	0	0	15	41
Spruce Fork	2	8	3	18	2	14	1	5	0	0	8	45
Clear Fork	0	0	1	8	3	12	3	12	1	7	8	39
Twentymile Creek	7	32	15	71	0	0	0	0	0	0	22	103
Twelvepole Creek	4	12	0	0	0	0	0	0	0	0	4	12
Total	23	76	28	137	7	38	6	19	2	12	66	282

[1]Because there were only two Mined/Residential sites, this EIS class was not used in any of the analyses for this report.

The samples taken from the Twelvepole Creek Watershed (four Unmined EIS class sites) were made up of a mix of D-frame kick net and Surber sampler data that were inseparable by sampler type. Therefore, these data could not be standardized and were removed from the EIS analysis for the D-frame kick net data set.

These data reduction procedures lowered the total number of D-frame kick net samples for EIS analysis from 282 (Table 3-2) to 215 (Table 3-4). The U.S. EPA Region 3 collected 150 (69.8%) of these samples and the other organizations collected 65 (30.2%) of these samples. Hence, these other organizations provided 43% more samples for analysis than the U.S. EPA Region 3 had collected. These samples also provided information from 23 additional sites in the Unmined, Filled, Filled/Residential, and Mined EIS classes. However, these additional samples were not distributed evenly across watersheds and EIS classes. Only the U.S. EPA Region 3 collected data from the Mud River, Spruce Fork, and Clear Fork Watersheds and the majority (85%) of the samples collected by the private organizations were collected from the Twentymile Creek Watershed. As a result, the additional data provided by the private organizations were skewed to conditions in the Twentymile Creek Watershed, especially for sites in the Filled EIS class. Furthermore, 100% of the data collected by the private organizations during autumn 2000 and winter 2001 were collected from the Twentymile Creek Watershed. Therefore, comparisons made using data that were collected during these two seasons do not represent conditions across the entire study area, and have less than half the number of samples that were collected during the other seasons.

Table 3-3. Correlation and significance values for the duplicate samples collected by the

U.S. EPA Region 3 with the WVSCI and standardized WVSCI metrics.

Metric	R	p-value
Total Number of Families Rarefied to 100 individuals	0.863	<0.001
Total Number of Ephemeroptera, Plecoptera, and Trichoptera (EPT) Families Rarefied to 100 individuals	0.897	<0.001
WVSCI Rarefied to 100 individuals	0.945	<0.001

Table 3-4. Number of sites and D-frame kick net samples used for comparing EIS classes after the data set had been reduced.

Watershed		EIS Class								Total	
		Unmined		Filled		Filled/ Residential		Mined			
		Site	Samp	Site	Samp	Site	Samp	Site	Samp	Site	Samp
Mud River	U.S. EPA	3	9	3	15	1	5	1	1	8	30
	Private	0	0	0	0	0	0	0	0	0	0
Island Creek	U.S. EPA	3	7	4	15	1	5	0	0	8	27
	Private	4	6	2	3	0	0	1	1	7	10
Spruce Fork	U.S. EPA	2	7	3	13	2	10	1	5	8	35
	Private	0	0	0	0	0	0	0	0	0	0
Clear Fork	U.S. EPA	0	0	1	5	3	10	3	9	7	24
	Private	0	0	0	0	0	0	0	0	0	0
Twenty-mile Creek	U.S. EPA	2	9	5	25	0	0	0	0	7	34
	Private	6	18	10	37	0	0	0	0	16	55
Total	U.S. EPA	10	32	16	73	7	30	6	15	38	150
	Private	10	24	12	40	0	0	1	1	23	65

3.2. Data Quality Assurance/Quality Control

The biological, water chemistry, and habitat data were received in a variety of formats. Data were exported from their original formats into the Ecological Data Application System (EDAS), a customized relational database application (Tetra Tech, Inc., 1999). The EDAS allows data to be aggregated and analyzed by customizing the pre-designed queries to calculate a variety of biological metrics and indices.

Throughout the process of exporting data, the original data sources were consulted for any questions or discrepancies that arose. First, the original electronic data files were consulted and proofread to ensure that the data had been migrated correctly from the original format into the EDAS database program. If the conflict could not be resolved in this manner, hard copies of data reports were consulted, or, as necessary, the mining companies and/or the organizations who had originally provided the data were consulted. As data were migrated, Quality Assurance/Quality Control (QA/QC) queries were used to check for import errors. If any mistakes were discovered as a result of one of these QA/QC queries, the entire batch was deleted, re-imported, and re-checked. After all the data from a given source had been migrated, a query was created which duplicated the original presentation of the data. This query was used to check for data manipulation errors. Ten percent of the original samples were checked at random. If the data failed this QC check, they were entirely deleted, re-imported, and subjected to the same QC routine until they were 100% correct.

The EDAS contained separate Master Taxa tables for fish and benthic macroinvertebrates. Both Master Taxa tables contained a unique record for each taxonomic name, along with its associated ecological characteristics (i.e., preferred habitat, tolerance to pollution). To ensure consistency, Master Taxa lists were generated from all of the imported MTM/VF data. Taxonomic names were checked against expert sources, such as Merritt and Cummins (1996), Robins et al. (1991) and the online taxonomic database, Integrated Taxonomic Information System (ITIS, www.itis.usda.gov). Discrepancies and variations in spellings of taxonomic names were identified and corrected in all associated samples. Any obsolete scientific names were updated to the current naming convention to ensure consistency among all the data. Each taxon's associated ecological characteristics were also verified to assure QC for biological metrics generated from that ecological information. Different organizations provided data at different levels of taxonomic resolution. Because the WVSCI utilizes benthic information at the Family level, the benthic macroinvertebrate Master Taxa table was used to collapse all of the data to the Family level for consistency in analysis.

Minimum Detection Limits (MDLs) represent the smallest amount of an analyte that can be detected by a given chemical analysis method. While some methods are very sensitive and, therefore, can detect very small quantities of a particular analyte, other methods are less sensitive and have higher MDLs. When an analytical laboratory is unable to detect an analyte, the value is reported as "Below Detection", and the MDL is given. For the purpose of statistical analysis, the "Below Detection" values were converted to ½ of the methods' MDLs.

3.3. Summary of Analyses

The fish database and the macroinvertebrate database were analyzed separately to: 1) determine if the biological condition of streams in areas with MTM/VF operations is degraded relative to the condition of streams in unmined areas and 2) determine if there are additive biological impacts to streams where multiple valley fills are located. The statistical approach to evaluate these two objectives was the same for fish and macroinvertebrates. To address the first objective, EIS classes (Filled, Filled/Residence, Mined, and Unmined) were compared using one-way analysis of variance (ANOVA). Assumptions for normality and equal variance were assessed using the Shapiro-Wilk Test for normality and Brown and Forsythe's Test for homogeneity of variance. If necessary, transformations were applied to the data to achieve normality and/or stabilize the variance. Significant differences ($p < 0.05$) among EIS classes were followed by the Least Square (LS) Means procedure using Dunnett's adjustment for multiple comparisons to test whether the Filled, Filled/Residence, and Mined EIS classes were significantly different ($p < 0.01$) from the Unmined EIS class. Additive sites from two watersheds were analyzed to evaluate the second objective. Trends in biological condition along the mainstem of Twentymile Creek and Twelvepole Creek were examined using Pearson correlations and regression analysis. Pearson correlations were also used to investigate correlations between biological endpoints and water chemistry parameters. Box plots were generated to display the data across EIS classes and scatter plots were created to show relationships between biological endpoints and chemistry parameters.

3.3.1. Summary of Fish Analysis

Endpoints for the fish analysis were the site averages for the Mid-Atlantic IBI and the site averages for the nine individual metrics that comprise the IBI (Table 1-2). Site averages were used in the analysis since the number of samples taken at a site was inconsistent across sites. Some study sites had been sampled only once, and there were also sites in the database that had been sampled on two or three separate occasions. Mean IBI and component metric values were calculated for all sites sampled multiple times. The mean values were used in all subsequent analyses. Figure 3-1 shows that there was no consistent difference between seasons or years, although there was scatter among observations at some sites. Log-transformed site (geometric) mean chemical concentrations were used as the endpoints for the chemistry analysis.

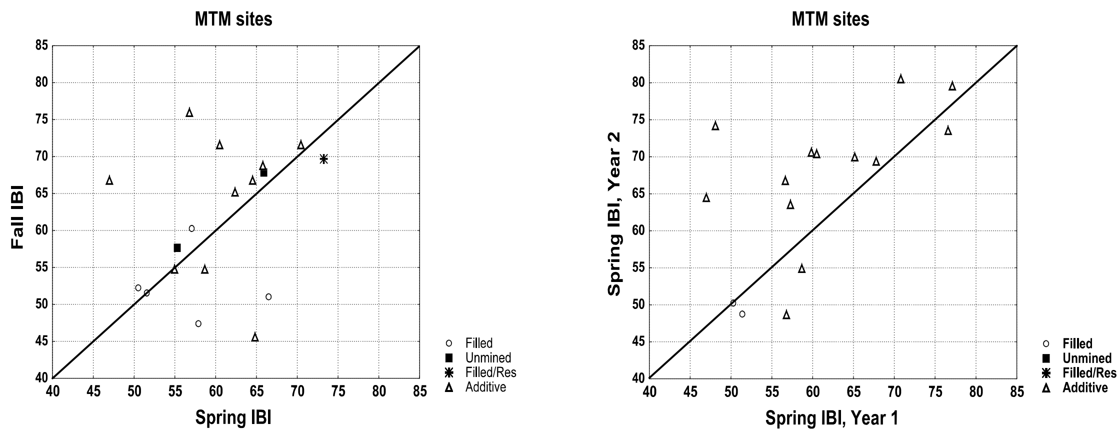


Figure 3-1. Scatter plots showing IBI scores of sites sampled multiple times. The left plot shows autumn samples versus spring samples and the right plot shows spring Year 2 samples versus spring Year 1 samples.

3.3.2. Summary of Macroinvertebrate Analysis

Endpoints for the macroinvertebrate analysis were the WV SCI and its component metrics (Total taxa richness, Ephemeroptera-Plecoptera-Trichoptera [EPT] taxa richness, Hilsenhoff Biotic Index [HBI], % dominant 2 taxa, % EPT abundance, and % Chironomidae abundance). Richness metrics and the WV SCI were rarefacted to 100 organisms to adjust for sampling effort. Comparisons among EIS classes were made for each season (Spring 1999 [April to June], Autumn 1999 [October to December], Winter 2000 [January to March], Spring 2000, Autumn 2000, and Winter 2001). Data for Summer 1999 (July to September) were not compared because of a lack of samples (n= 2) for the Unmined EIS class (i.e., the relative control). Furthermore, in some seasons there were insufficient samples (n < 3) for the Mined and Filled/Residence classes. The WVSCI scores were correlated against key water quality parameters using mean values for each site. Only water chemistry data that were collected at or close to the time of benthos sample collection were used in this analysis.

Habitat data was not evaluated due to the fact that it was not collected consistently and in many cases was collected only once at a site.

4. RESULTS

4.1. Fish Results

4.1.1. IBI Calculation and Calibration

Generally, larger watersheds tend to be more diverse than smaller watersheds (i.e., Karr et al. 1986, Yoder and Rankin 1995). This was found to be true in the MTM/VF study where the smallest headwater streams often had either no fish present or only one or two species present and the large streams had 15 to 27 fish species present (Figure 4-1). To ensure that differences among fish communities were due to differences in stream health and not from the natural effect of watershed size, the three richness metrics (i.e., Native Intolerant Taxa, Native Cyprinidae Taxa and Native Benthic Invertivores) from the Mid-Atlantic Highlands IBI (Section 1.5) were standardized to a 100-km^2 watershed. If the calibration was correct, then there should have been no residual relationship between catchment area and IBI scores. The resultant IBI scores were plotted against catchment area (Figure 4-2) which showed that there was no relationship.

The Mid-Atlantic IBI was not calculated if the catchment area was less than 2.0 km^2. If fewer than ten fish were captured in a sample, then the IBI was set to zero (McCormick et al. 2001). This occurred in six samples. All six of these samples were in relatively small catchments (i.e., 2.0 to 5.0 km^2), where small samples are likely (Figure 4-2). Because small samples may be due to natural factors, these samples were excluded from subsequent analysis..

4.1.2. IBI Scores in EIS Classes

The distributions of IBI scores in each of the EIS classes are shown in Figure 4-3. Distributions of the nine component metrics of the IBI are shown in Appendix B. For comparison, the regional reference sites sampled by the PSU in Big Ugly Creek were also plotted. Figure 4-3 shows that the Filled and Mined classes have lower overall IBI scores than the other EIS classes. The Filled/Residential class had higher IBI scores than any other class. The Filled/Residential class and the Unmined class had median scores that were similar to the regional reference sites. Figure 4-3 shows that more than 50% of the Filled and Mined sites scored “poor” according to the ratings developed by McCormick et al. (2001). Unmined and regional reference sites were primarily in the “fair” range and Filled/Residential sites were mostly in the “good” ranges.

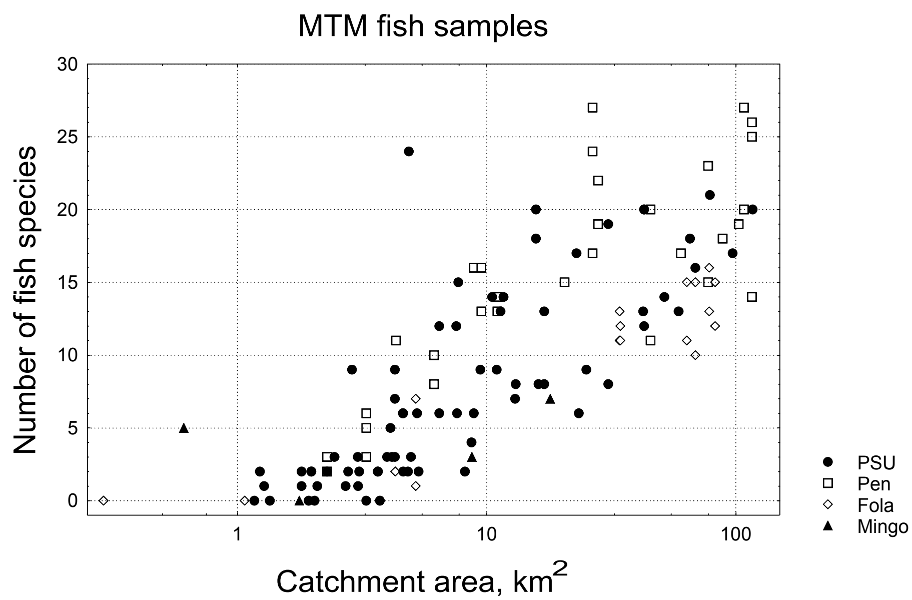


Figure 4-1. Number of fish species captured versus stream catchment area. Symbols identify sampling organizations: PSU=Penn State; Pen = Pen Coal (REIC); Fola = Fola Coal (Potesta); Mingo = Mingo-Logan Coal (BMI).

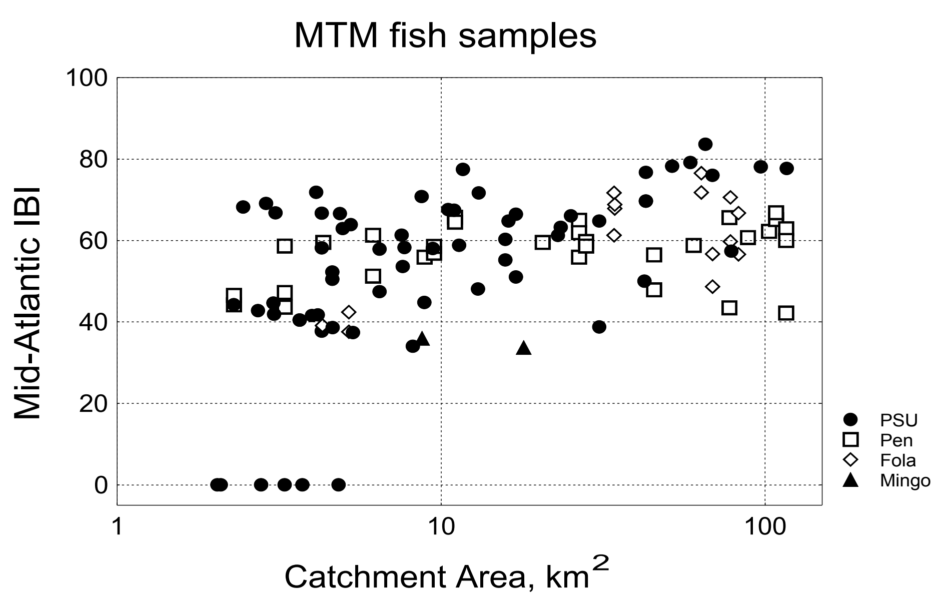


Figure 4-2. Calculated Fish IBI and watershed catchment area, all MTM fish samples from sites with catchment > 2km^2. Symbols identify sampling organizations: PSU=Penn State; Pen = Pen Coal (REIC); Fola = Fola Coal (Potesta); Mingo = Mingo-Logan Coal (BMI).

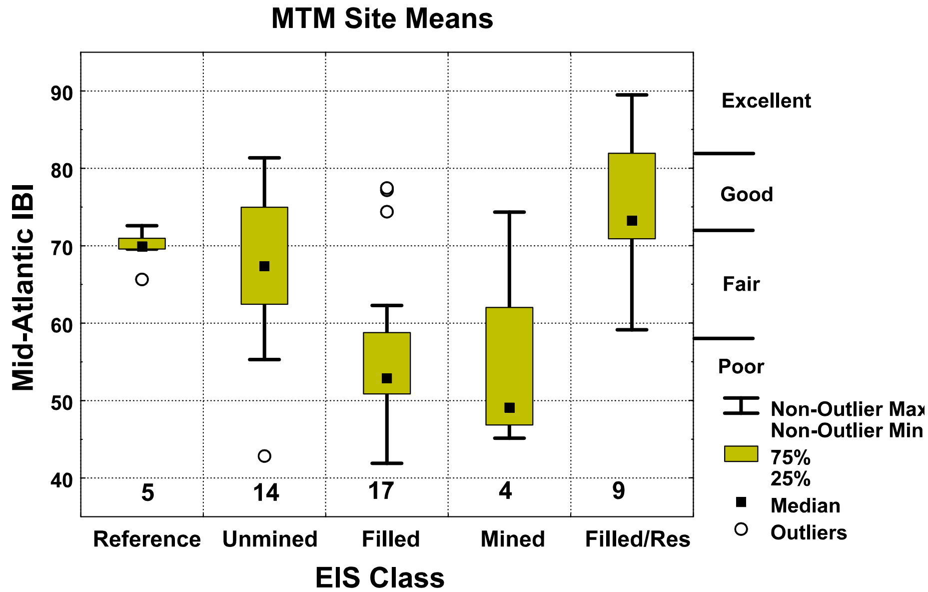


Figure 4-3. A Box-and-Whisker plot of the mean IBI scores from sampling sites in five EIS classes. Catchments less than 2 km^2 and samples with less than ten fish were excluded. Numbers below boxes indicate sample size. Reference sites were the five regional reference sites in Big Ugly Creek, outside of study area. All other sites were in the MTM study area. Assessment categories (McCormick et al.2001) are shown on right side.

A one-way ANOVA was used to test for differences among EIS classes and the LS Means procedure with Dunnett's adjustment was used to compare each class to the Unmined class. The ANOVA showed that differences among the EIS classes were statistically significant (Table 4-1) and the LS Means test showed that the IBI scores from the Filled sites were significantly lower than the IBI scores from the Unmined sites (Table 4-2). The Filled/ Residential class had higher IBI scores than the Unmined sites (Figure 4-3). The IBI scores from Mined sites were lower than the IBI scores from Unmined sites. However, the difference was only marginally significant. This is most likely due to the small sample of Mined sites (n=4). Diagnostics on the IBI analysis indicated that variance was homogeneous and residuals of the model were normally distributed (Figure 4-4 and Appendix B).

The individual metrics that comprise the IBI are not uniform in their response to stressors (McCormick et al. 2001). While some metrics may respond to habitat degradation, other metrics may respond to organic pollution or toxic chemical contamination. Of the nine metrics in the IBI, two (i.e., the number of cyprinid species and the number of benthic invertivore species) were significantly different among the EIS classes. (Appendix B). On average, Filled sites were missing one species of each of these two groups compared to Unmined sites. The third taxa richness metric, Number of Intolerant Species, was not different between Filled and Unmined sites (Appendix B). One additional metric, Percent Tolerant Individuals, showed increased degradation in Filled and Mined sites compared to Unmined sites, on average, but the difference was not statistically significant (Appendix B). Four metrics, Percent Cottidae, Percent Gravel Spawners, Percent Alien Fish and Percent Large Omnivores, were dominated by zero values (Appendix B). Because of the zero values and the resultant non-normal distribution, parametric hypothesis tests would be problematic.

It was concluded from this analysis that the primary causes of reduced IBI values in Filled sites were reductions in the number of minnow species and the number of benthic invertivore species. These two groups of fish are dominant in healthy Appalachian streams. Secondary causes of the reduction of IBI scores in Filled sites are decreased numbers of intolerant taxa, and increased percentages of fish tolerant to pollution. Although Filled sites had IBI scores that were significantly lower than Unmined sites (Table 4-3), several Filled and Mined sites had relatively high IBI scores, similar to regional reference and Unmined sites. In addition, the Filled/Residential sites had higher overall IBI scores. Field crews had observed that there were very few or no residences in the small watersheds of the headwater stream areas. This suggests that the sites where fills and residences were co-located occurred most frequently in larger watersheds and that watershed size may buffer the effects of fills and mines. This possibility was examined and it was found that Filled, Mined, and Filled/Residential sites in watersheds with areas greater than 10 km^2 had fair to good IBI scores. However, Filled and Mined sites in watersheds with areas less than 10 km^2 often had poor IBI scores (Figure 4-5A). Of the 14 sites in watersheds with areas greater than 10 km^2, four were rated fair and ten were rated good or better (Figure 4-5A). Of the 17 sites in watersheds with areas less than 10 km^2, only three rated fair and 14 rated poor (Figure 4-5). In contrast, the control and reference sites showed no overall association with catchment area (Figure 4-5B). The smallest sites (i.e., watershed areas < 3.0 km^2) were highly variable, with three of the five smallest sites scoring poor.

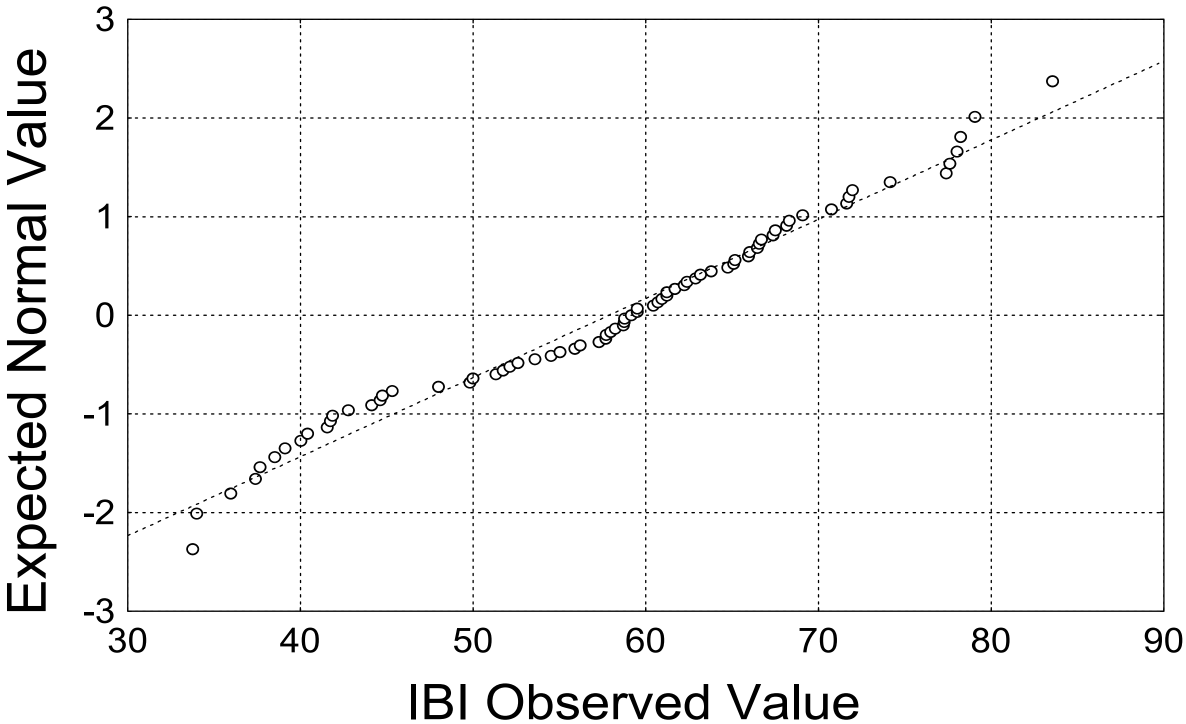


Figure 4-4. Normal probability plot of IBI scores from EIS classes.

Table 4-1. The ANOVA for IBI scores among EIS classes (Unmined, Filled, Mined, and Filled/Residential).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2335.56	778.52	6.70	0.0009
Error	40	4651.31	116.28		
Corrected Total	43	6986.87			

R-Square	Coefficient of Variance	Root MSE	Index Mean
0.334	17.022	10.783	63.350

Table 4-2. Dunnett's test comparing IBI values of EIS classes to the Unmined class, with the alternative hypothesis that IBI < Unmined IBI (one-tailed test).

EIS Class	N	Mean	Standard Deviation	Dunnett’s P-Value
Filled	**17**	**56.8**	**10.6**	**0.0212**
Filled/Residential	**9**	**74.6**	**10.7**	**0.9975**
Mined	**4**	**54.4**	**13.4**	**0.0685**
Unmined	**14**	**66.7**	**10.3**	**--**

The effect of fills was statistically stronger in watersheds with areas less than 10 km^2 (Table 4-3). Filled sites had an average of one fewer Cyprinidae species, 1.6 fewer benthic invertivore species, 20% more tolerant individuals, and a mean IBI score that is 14 points lower than Unmined sites (Table 4-3). In addition, Intolerant Taxa, % Cottidae and % Gravel Spawners decreased slightly in the filled sites and the % Macro Omnivores increased slightly (Table 4-3). There were too few small Mined sites (n=3) and too few small Filled/Residential sites (n=2) to test against the Unmined sites within the small size category.

There is no definitive test to determine whether the high IBI scores of the Filled/Residential sites in this data set are due solely to large catchment areas or if there may be other contributing factors. The Filled/Residential class is consistent with the relationship observed in the Filled sites, that large catchments are less susceptible to the effects of fills and mines. A definitive test could be conducted if data were collected from several small Filled/Residential catchments.

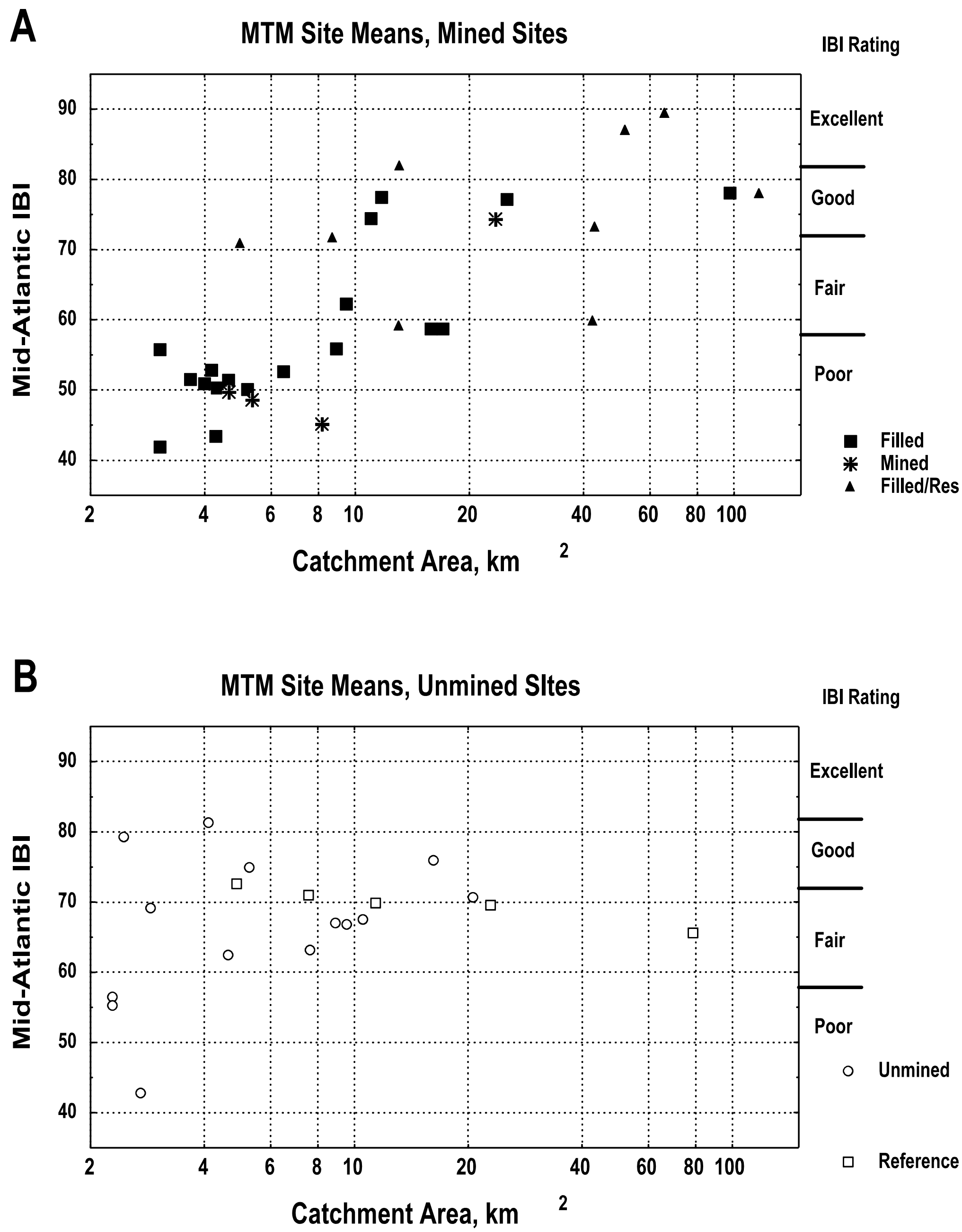


Figure 4-5. The IBI scores for different site classes, by watershed area. Assessment categories (McCormick et al.2001) are shown on right. A) Filled, Mined, and Filled/ Residential sites. B) Unmined and Reference (Big Ugly Creek) sites.

Table 4-3. The results of t-tests of site mean metric values and the IBI in Unmined and Filled sites in watersheds with areas less than 10 km^2 (N = 11 Unmined, N = 12 Filled).

	Mean Unmined	Mean Filled	t-value	p
Cyprinidae Taxa	5.41	4.37	2.93	0.008
Intolerant Taxa	1.03	0.85	1.23	0.232
Benthic Invertivore Taxa	5.80	4.22	3.73	0.001
% Exotic	0.3	0.9	-0.65	0.524
% Cottidae	3.8	0.4	1.42	0.172
% Gravel Spawners	17.2	7.0	0.999	0.329
% Piscivore/Invertivores	34.8	38.8	-0.34	0.739
% Tolerant	71.8	93.8	-2.60	0.0167
% Macro Omnivore	1.4	4.8	-1.54	0.139
IBI	65.4	51.5	3.80	0.001

4.1.3. Additive Analysis

Sites on the mainstem of Twentymile Creek and all mining-affected sites in the Twelvepole Creek watershed have been identified as Additive sites, and were not included in the analysis of the EIS classes reported above. Instead, these sites were considered to be subject to multiple and possibly cumulative sources (i.e., VFs, historic mining, non-point runoff, untreated domestic sewage, non-permitted discharges).

The Twelvepole Creek watershed, in particular, has mixed land uses and has several mining techniques in use. The stream valleys are often populated with residences and livestock. Mining in the Twelvepole watershed includes deep mining, contour mining, and mountaintop removal/VF. In contrast, there is little or no residential land use in the Twentymile Creek watershed and all human activities in the Twentymile Creek are related to mining (i.e., logging and grubbing).

The IBI scores of sites in three streams (i.e., Kiah Creek, Trough Fork, and Twelvepole Creek) in the Twelvepole Creek Watershed are shown in Figure 4-6. Most of the sites are scored in the “fair” range, although a few observations extend into the “good” and “poor” ranges (Figure 4-6). There is no apparent pattern in these scores and there are no trends from upstream to downstream in either of the larger streams (i.e., Kiah Creek and Twelvepole Creek).

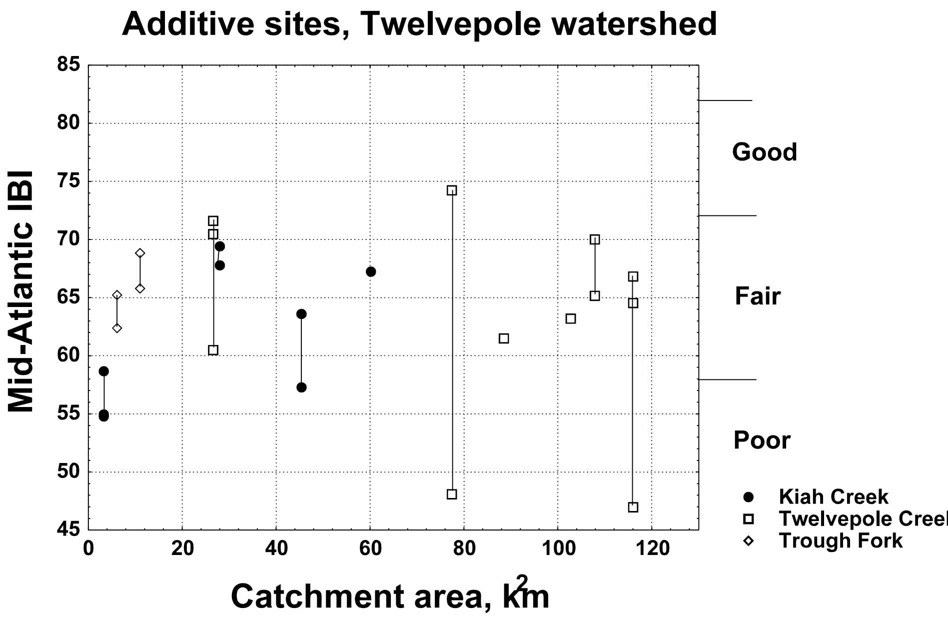
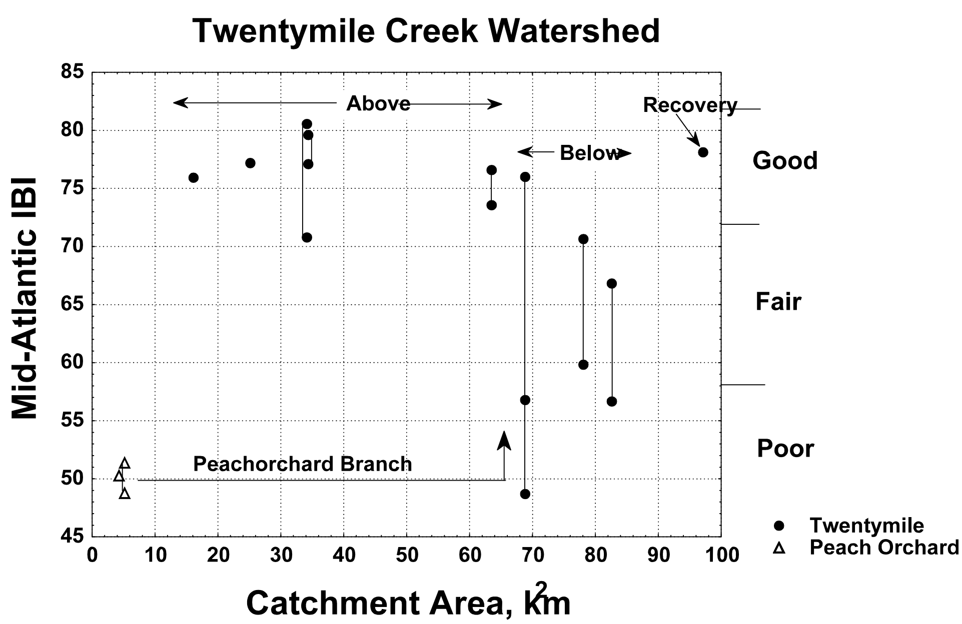


Figure 4-6. The IBI scores from the additive sites in the Twelvepole Creek Watershed. Multiple observations from single sites are connected with a vertical line.

Figure 4-7. IBI scores from additive sites and Peachorchard Branch in the Twentymile



Creek Watershed. Multiple observations from single sites are connected with a vertical line.

Overall, the IBI scores in the Twentymile Creek watershed were higher than those in Twelvepole Creek. There was a trend, from upstream to downstream, among the scores from the Twentymile Creek Watershed (Figure 4-7). Above Peachorchard Branch, which has a catchment area smaller than 68 km^2, sites on the mainstem of Twentymile Creek were uniformly in the "good" range of IBI scores, with moderate variability. Below the confluence of Peachorchard Branch, IBI scores decrease overall and are more variable (Figure 4-7). Farther downstream (i.e., Site PSU.54), the IBI score was higher (i.e., 78), indicating potential recovery from the stressors in the lower portion of the stream. With a range of 48 to 52, Peachorchard Branch had among the lowest IBI scores in the Twentymile Creek Watershed.

4.1.4. Associations With Potential Causal Factors

The correlations between IBI scores and water quality parameters that are potential stressors (i.e., DO, pH, nutrients, TDS, TSS, salts, and metal concentrations) were examined. For the correlation analysis, site mean IBI scores and log-transformed site (geometric) mean chemical concentrations were used. The correlation analysis was restricted to sites in watersheds with areas smaller than 10.0 km^2. The IBI scores decreased with the increased concentrations of several water quality parameters, and decreased significantly with increased zinc and sodium (Table 4-4). However, these correlations do not imply causal relationships between water quality parameters and fish community condition. Other substances or processes associated with mining activity (i.e., erosion, sedimentation), but not measured, could also be proximal causal factors.

Table 4-4. Pearson correlations among the site means of selected water quality measurements and IBI scores, including all sites in watersheds with areas smaller than 10 km^2.

	Log Cr	Log Mg	Log Ni	Log	Log Na	Log SO_4	Log TDS	Log Zn
Log Mg	0.11							
Log Ni	-0.08	0.53						
Log (NO_3+NO_2)	0.40	0.65	0.37					
Log Na	0.16	0.40	-0.08	0.65				
Log SO_4	0.17	0.96	0.43	0.76	0.58			
Log TDS	0.27	0.42	-0.35	0.79	0.90	0.65		
Log Zn	0.50	0.34	0.12	0.47	0.34	0.38	0.42	
IBI	-0.35	-0.42	-0.33	-0.42	-0.60	-0.51	-0.47	-0.54

4.2. Macroinvertebrate Results

4.2.1. Analysis of Differences in EIS Classes

For each season, analyses were conducted to determine if there were any differences among the EIS classes. Only Unmined, Filled, Mined and Filled/Residential sites were used for these analyses. Analysis endpoints were the WVSCI and it's component metrics.

4.2.1.1. Spring 1999

This comparison only used U.S. EPA Region 3 data for each watershed. All of the tested metrics were significantly different among EIS classes using ANOVA, and each met the assumptions for normality and equal variance (Table 4-5). The WVSCI and the taxa richness metrics differed significantly between Unmined sites and both Filled and Filled/Residential sites in the LS Means test. Percent EPT Abundance was also significantly different between Unmined sites and Filled/Residential sites. Box plots for each metric comparison are in Appendix C.

4.2.1.2. Autumn 1999

This comparison used data collected by both the U.S. EPA Region 3 and the private organizations for each watershed. Only the WVSCI, Percent EPT and Percent Chironomidae Abundance were significantly different among EIS classes (Table 4-6). However, the Unmined sites were not significantly different from the other classes for these metrics. Box plots for each metric comparison are in Appendix C. Drought conditions occurred during this season, and streams were further impacted by a severe drought during the preceding summer.

Table 4-5. Results from ANOVA for benthic macroinvertebrates in spring 1999. Uses Unmined sites as a relative control for LS Means test. Total n = 34; Unmined n = 9, Mined n = 4, Filled n = 15, Filled/Residential n = 6.

Metric	p-value	Normality?	Equal Variance?	LS Means
WVSCI (Rarefied to 100 Organisms)	<0.0001	Yes	Yes	Filled and Filled/Residential
Total Taxa (Rarefied to 100 Organisms)	0.0001	Yes	Yes	Filled and Filled/Residential
EPT Taxa (Rarefied to 100 Organisms)	<0.0001	Yes	Yes	Filled and Filled/Residential
HBI	0.0017	Yes	Yes	
Percent Dominant Two Taxa (Arcsine Transformed)	0.0010	Yes	Yes	
Percent EPT Abundance (Arcsine Transformed)	0.0010	Yes	Yes	Filled/Residential
Percent Chironomidae Abundance (Arcsine Transformed)	0.0326	Yes	Yes	

Table 4-6. Results from ANOVA for benthic macroinvertebrates in autumn 1999. Uses Unmined sites as a relative control for LS Means test. Total n = 35, Unmined n = 6, Filled n = 23, Filled/Residence n = 6.

Metric	p-value	Normality?	Equal Variance?	LS Means
WVSCI (Rarefied to 100 Organisms)	0.0454	Yes	Yes	
Total Taxa (Rarefied to 100 Organisms)	0.3744	Yes	Yes	
EPT Taxa (Rarefied to 100 Organisms)	0.2401	Yes	Yes	
HBI	0.1299	Yes	Yes	
Percent Dominant Two Taxa (Arcsine Transformed)	0.2672	Yes	Yes	
Percent EPT Abundance (Arcsine Transformed)	0.0178	Yes	Yes	
Percent Chironomidae Abundance (Arcsine Transformed)	0.0253	Yes	Yes	

4.2.1.3. Winter 2000

This comparison used data collected by both the U.S. EPA Region 3 and the private organizations for each watershed. All of the tested metrics were significantly different among EIS classes, and each met the assumptions for normality (Table 4-7). The WVSCI and the HBI failed the test for equal variance. The WVSCI and the Total Taxa metrics differed significantly between Unmined sites and both Filled and Filled/Residential sites in the LS Means test. Percent EPT abundance was also significantly different between Unmined sites and Filled/Residential sites. Box plots for each metric comparison are in Appendix C.

4.2.1.4. Spring 2000

This comparison used only the data collected by the U.S. EPA Region 3 for each watershed. All of the tested metrics were significantly different among EIS classes, and each met the assumptions for normality (Table 4-8). The WVSCI, EPT Taxa, HBI, and Percent EPT Abundance failed the test for equal variance. The WVSCI and the taxa richness metrics differed significantly between Unmined sites and both Filled and Filled/Residence sites in the LS Means test. Percent EPT abundance in the Unmined sites was also significantly different than in Filled/Residence sites. Box plots for each metric comparison are in Appendix C.

4.2.1.5. Autumn 2000

This comparison used only the data collected by the private organizations for the Twentymile Creek watershed. No metrics were significantly different among EIS classes (Table 4-9). Box plots for each metric comparison are in Appendix C.

4.2.1.6. Winter 2001

This comparison used only the data collected by the private organizations for the Twentymile Creek watershed. The WVSCI, Total Taxa, EPT Taxa, and Percent Dominant 2 Taxa were significantly different among EIS classes (Table 4-10). The Unmined sites were significantly different than the Filled classes for the WVSCI and EPT Taxa, although both metrics failed the equal variance test. Box plots for each metric comparison are in Appendix C.

Table 4-7. Results from ANOVA for benthic macroinvertebrates in winter 2000. Uses Unmined sites as a relative control for LS Means test. Total n = 53, Unmined n = 18, Mined n = 4, Filled n =25, Filled/Residential n = 6.

Metric	p-value	Normality?	Equal Variance?	LS Means
WVSCI (Rarefied to 100 Organisms)	<0.0001	Yes	No	Filled and Filled/Residential
Total Taxa (Rarefied to 100 Organisms)	<0.0001	Yes	Yes	Filled and Filled/Residential
EPT Taxa (Rarefied to 100 Organisms)	<0.0001	Yes	Yes	Filled and Filled/Residential
HBI	<0.0001	Yes	No	
Percent Dominant Two Taxa (Arcsine Transformed)	<0.0001	Yes	Yes	
Percent EPT Abundance (Arcsine Transformed)	<0.0001	Yes	Yes	Filled and Filled/Residential
Percent Chironomidae Abundance (Arcsine Transformed)	<0.0001	Yes	Yes	

Table 4-8. Results from ANOVA for benthic macroinvertebrates in spring 2000. Uses Unmined sites as a relative control for LS Means test. Total n = 35, Unmined n = 10, Mined n = 5, Filled n = 15, Filled/Residence n = 5.

Metric	p-value	Normality?	Equal Variance?	LS Means
WVSCI (Rarefied to 100 Organisms)	0.0001	Yes	No	Filled and Filled/Residential
Total Taxa (Rarefied to 100 Organisms)	0.0004	Yes	Yes	Filled and Filled/Residential
EPT Taxa (Rarefied to 100 Organisms)	<0.0001	Yes	No	Filled and Filled/Residential
HBI	0.0002	Yes	No	
Percent Dominant Two Taxa (Arcsine Transformed)	<0.0001	Yes	Yes	
Percent EPT Abundance (Arcsine Transformed)	0.0027	Yes	No	Filled/Residential
Percent Chironomidae Abundance (Arcsine Transformed)	0.0020	Yes	Yes	

Table 4-9. Results from ANOVA for benthic macroinvertebrates in autumn 2000. Uses Unmined sites as a relative control for LS Means test. Total n = 15; Unmined n = 5, Filled n = 10.

Metric	p-value	Normality?	Equal Variance?	LS Means
WVSCI (Rarefied to 100 Organisms)	0.1945	Yes	Yes	
Total Taxa (Rarefied to 100 Organisms)	0.4744	Yes	Yes	
EPT Taxa (Rarefied to 100 Organisms)	0.1897	Yes	Yes	
HBI	0.7243	Yes	Yes	
Percent Dominant Two Taxa (Arcsine Transformed)	0.0846	Yes	Yes	
Percent EPT Abundance (Arcsine Transformed)	0.3200	Yes	Yes	
Percent Chironomidae Abundance (Arcsine Transformed)	0.4417	Yes	Yes	

Table 4-10. Results from ANOVA for benthic macroinvertebrates in winter 2001. Uses Unmined sites as a relative control for LS Means test. Total n = 16, Unmined n = 6, Filled n = 10.

Metric	p-value	Normality?	Equal Variance?	LS Means
WVSCI (Rarefied to 100 Organisms)	0.0110	Yes	No	Filled
Total Taxa (Rarefied to 100 Organisms)	0.0275	Yes	Yes	
EPT Taxa (Rarefied to 100 Organisms)	0.0074	Yes	No	Filled
HBI	0.4874	Yes	Yes	
Percent Dominant Two Taxa (Arcsine Transformed)	0.0012	Yes	Yes	
Percent EPT Abundance (Arcsine Transformed)	0.3449	Yes	Yes	
Percent Chironomidae Abundance (Arcsine Transformed)	0.1180	Yes	Yes	

4.2.2. Evaluation of Twentymile Creek

Box plots were used to compare benthic macroinvertebrate metrics in the major watersheds during spring 1999, autumn 1999, winter 2000, and spring 2000. Only data from Twentymile Creek was available for autumn 2000 and winter 2001 and it was necessary to examine whether the EIS data collected from the Twentymile Creek Watershed was similar to the EIS data collected from the other four watersheds. Clear Fork could not be used in this watershed analysis, since data for Clear Fork were limited (i.e., there were no Unmined sites and only one Filled site).

No consistent differences in the benthic metrics between the Unmined sites and among watersheds were observed (Appendix C). In contrast, there were consistent differences in the benthic metrics between Filled sites and among watersheds in each season except autumn 1999. Total Taxa, EPT Taxa, Percent EPT Abundance, and the WVSCI were consistently better in Twentymile Creek and Island Creek watersheds than in the Mud River and Spruce Fork watersheds (Appendix C).

4.2.3. Macroinvertebrate and Water Chemistry Associations

The WVSCI scores were correlated against key water quality parameters using mean values for each site. Only water chemistry data that were collected at or close to the time of benthos sample collection were used in this analysis.

The strongest associations were negative correlations between the WVSCI and measures of individual and combined ions (Table 4-11, Appendix D). The WVSCI was also negatively correlated with the metals Beryllium, Selenium, and Zinc.

4.2.4. The Effect of Catchment Area on the WVSCI

The WVSCI and its component metrics had not been evaluated for potential effects related to stream size because of a lack of catchment area data during the original index development. The WVSCI and its component metric scores calculated from the MTM/VF data were plotted against catchment area. A Pearson correlation analysis was also run on these data to investigate whether stream size influenced these scores for the MTM/VF EIS analysis. This analysis was only conducted for the sites in the Unmined EIS class in order to limit any confounding variation due to anthropogenic sources.

There were 20 Unmined sites available for this analysis. However, one site was dropped because catchment area data for that site was unavailable. Because sample size varied greatly

Table 4-11. Results from Pearson correlation analyses between the WVSCI rarefied to 100 organisms and key water quality parameters.

Parameter	n	R	P-value
Alkalinity	53	-0.660	<0.001
Total Aluminum	47	-0.208	0.161
Total Beryllium	52	-0.298	0.032
Total Calcium	53	-0.624	<0.001
Total Chromium	53	-0.043	0.761
Conductivity	53	-0.690	<0.001
Total Copper	53	-0.238	0.086
Hardness	23	-0.650	0.001
Total Iron	49	-0.189	0.193
Total Magnesium	53	-0.569	<0.001
Total Manganese	49	-0.241	0.095
Total Nickel	53	-0.166	0.235
Nitrate/Nitrite	21	-0.362	0.106
DO	60	0.031	0.815
Total Phosphorus	53	-0.165	0.237
Total Potassium	53	-0.527	<0.001
Total Selenium	51	-0.476	<0.001
Total Sodium	53	-0.572	<0.001
Sulfate	53	-0.598	<0.001
Total Dissolved Solids	53	-0.371	0.006
Total Zinc	53	-0.343	0.012

among seasons and was very low in some seasons (i.e., n = 5 or 6), the mean score for each site was used in the analyses.

Neither correlation analyses (Table 4-12) nor scatter plots (Figure 4-8) showed an effect of catchment area on the WVSCI and its metric scores. Analyses with arcsin transformed proportion metrics (i.e., Percent Dominant Two Taxa, Percent EPT Taxa, and Percent Chironomid Taxa) also showed no relationship to catchment area ® = 0.269, -0.144, and 0.090, respectively)

Although no relationship was found, these analyses were limited by the relatively low sample sizes available, and the limited range in catchment area (0.29 – 5.26 km^2) data for Unmined sites. Additional data for larger and relatively undisturbed stream sites within the MTM/VF footprint is necessary to examine stream size effects for the three larger (i.e., area > 40 km^2) Filled/Residence sites. It is unclear whether such sites exist in this area.

Table 4-12. Pearson correlation values and p-values for means of metric scores at Unmined sites (n = 19) versus catchment area.

Metric	R	p-value
Tot_S100	-0.157	0.520
EPT_S100	-0.165	0.501
HBI	0.228	0.348
Dom2Pct	0.255	0.293
EPTPct	-0.168	0.493
ChirPct	0.087	0.724
WVSCI100	-0.312	0.194

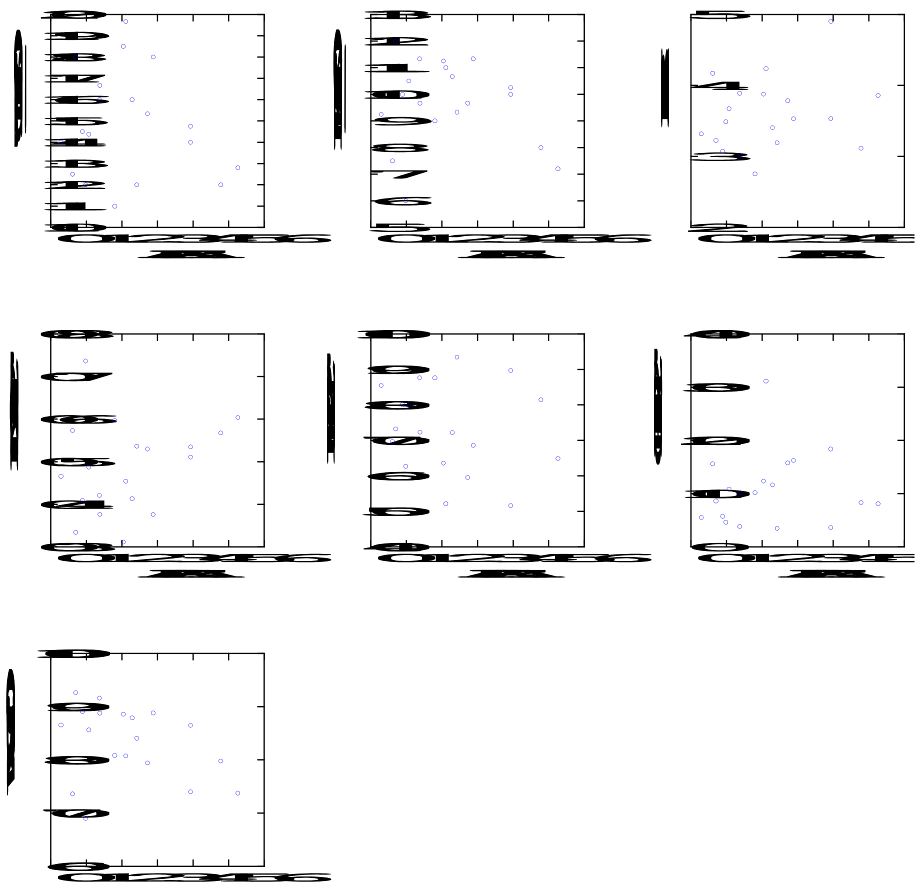


Figure 4-8. The WVSCI and its metric scores versus catchment area in Unmined streams.

4.2.5. Additive Analysis

Multiple sites on the mainstem of Twentymile Creek were identified as Additive sites and were included in an analysis to evaluate impacts of increased mining activities in the watershed across seasons and from upstream to downstream of the Twentymile Creek. Cumulative river kilometer was calculated for each site along Twentymile Creek as the distance from the uppermost site, Rader 8. The total distance upstream to downstream was approximately 17 kilometers. Sites were sampled during four seasons, Autumn 1999 (n = 19), Winter 2000 (n = 23), Autumn 2000 (n = 24) and Winter 2001 (n = 26). Pearson correlations between cumulative river kilometer and the WVSCI and it's component metrics were calculated for each season (Table 4-13). The number of metrics that showed significant correlations with distance along the mainstem increased across seasons. The WVSCI was significantly correlated with cumulative river kilometer in Winter 2000, Autumn 2000 and Winter 2001. In Winter 2001, four of the six individual metrics also showed significant correlations with distance along the mainstem of Twentymile Creek. A linear regression of the WVSCI with cumulative river kilometer indicated that the WVSCI decreased approximately one point upstream to downstream for every river kilometer (Table 4-14).

Table 4-13. Pearson correlation values and p-values for metric scores at Additive sites on Twentymile Creek versus cumulative river kilometer by season.

Metric	Autumn 1999	Winter 2000	Autumn 2000	Winter 2001
Tot_S100	-0.582 (0.009)	0.051 (0.8169)	-0.670 (<.001)	-0.462 (0.018)
EPT_S100	-0.480 (0.038)	-0.230 (0.196)	-0.688 (<.001)	-0.593 (0.002)
HBI	-0.210 (0.387)	-0.227 (0.296)	-0.228 (0.284)	0.410 (0.037)
Dom2Pct	0.360 (0.130)	0.521 (0.011)	0.626 (0.001)	0.545 (0.004)
EPTPct	0.018 (0.940)	-0.004 (0.986)	0.145 (0.499)	-0.235 (0.248)
ChirPct	-0.075 (0.759)	-0.377 (0.076)	-0.048 (0.824)	0.091 (0.658)
WVSCI100	-0.353 (0.138)	0.762 (<.001)	-0.627 (0.001)	-0.608 (0.001)

Table 4-14. The Regression for WVSCI versus Cumulative River Mile for Additive Sites in Twentymile Creek Winter 2001.

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	658.99	658.99	14.05	0.0010
Error	24	1125.55	46.90		
Corrected Total	25	1784.54			

R-Square	Coefficient of Variance	Root MSE	WVSCI Mean
0.369	8.27	6.848	82.80

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	92.66	2.95	31.38	<.0001
Cumulative River Km	-1.14	0.30	-3.75	0.001

5. DISCUSSION AND CONCLUSIONS

5.1. Fish Discussion and Conclusions

From the analysis of the fish data among the EIS classes, it was determined that IBI scores were significantly reduced in streams below VFs, compared to unmined streams, by an average of 10 points, indicating that fish communities were degraded below VFs. The IBI scores were similarly reduced in streams receiving drainage from historic mining or contour mining, compared to unmined streams. Nearly all filled and mined sites with catchment areas smaller than 10.0 km^2 had “poor” IBI scores, whereas filled and mined sites with catchment areas larger than 10.0 km^2 had “fair” or “good” IBI scores. In the small streams, IBI scores from Filled sites were an average of 14 points lower than the IBI scores from Unmined sites. Most Filled/Residential sites were in larger watersheds (i.e., areas > 10.0 km^2), and Filled/Residential sites had “fair” or “good” IBI scores.

From the additive analysis, it was determined that the Twelvepole Creek Watershed, in which the land use was mixed residential and mining, had “fair” IBI scores in most samples, and there are no apparent additive effects of the land uses in the downstream reaches of the watershed. Also, Twentymile Creek, which has only mining-related land uses, has “Good” IBI scores upstream of the confluence with Peachorchard Creek, and “Fair” and “Poor” scores for several miles downstream of the confluence with Peachorchard Creek tributary. Finally, Peachorchard Creek has “Poor” IBI scores, and may contribute contaminants or sediments to Twentymile Creek, causing degradation of the Twentymile IBI scores downstream of Peachorchard Creek.

5.2. Macroinvertebrate Discussion and Conclusions

The results of the macroinvertebrate analyses showed significant differences among EIS classes for the WVSCI and some of its component metrics in all seasons except autumn 2000. Differences in the WVSCI were primarily due to lower Total Taxa, especially for mayflies, stoneflies, and caddisflies, in the Filled and Filled/Residential EIS classes.

Sites in the Filled/Residential EIS class usually scored the worst of all EIS classes across all seasons (Appendix C). It was not determined why the Filled/Residential class scored worse than the Filled class alone. U.S. EPA (2001 Draft) found the highest concentrations of Na in the Filled/Residential EIS class, which may have negatively impacted these sites compared to those in the Filled class.

When the results for Filled and Unmined sites alone were examined, significant differences were observed in all seasons except autumn 1999 and autumn 2000. This can be seen in the plots of the WVSCI, Total Taxa, and EPT Taxa versus season (Figures 5-1, 5-2a and

5-2b). The lack of differences between Unmined and Filled sites in autumn 1999 was due to a decrease in Total Taxa and EPT Taxa in Unmined sites relative to a lack of change in Filled sites. These declines in taxa richness metrics in Unmined sites was likely a result of the drought conditions of the summer 1999, which caused more Unmined sites to go dry or experience severe declines in flow relative to Filled sites (Green et al., 2000). Wiley et al. (2001) also found that Filled sites have daily flows that are greater than those in Unmined sites during periods of low discharge. Despite the relatively drier conditions in Unmined sites during autumn 1999, WVSCI scores and EPT Taxa richness increased in later seasons to levels seen in the spring 1999 season whereas values for Filled sites stayed relatively low.

The lack of statistical differences between Unmined and Filled classes in the autumn 2000 appears to be due to a decline of Total Taxa richness in Unmined sites coupled with an increase in Total Taxa richness in Filled sites (Figures 5-1, 5-2 and 5-3). Filled sites had higher variability in WVSCI scores and metric values than did Unmined sites during the autumn 2000, which also contributed to the lack of significant differences. It is important to note that this comparison only uses data from the Twentymile Creek Watershed. Hence, the lack of differences in metrics during the autumn 2000 between Unmined and Filled sites is only relevant for the Twentymile Creek watershed, and not the entire MTM/VF study area examined in the preceding seasons. Similarly, data for winter 2001 is only representative of the Twentymile Creek watershed, but it is noteworthy that these data did show that Unmined and Filled sites were significantly different. It was also found that Filled sites in the Twentymile Creek Watershed scored better than filled sites in the Mud River and Spruce Fork Watersheds in all seasons except for autumn 1999. These differences among watersheds indicate biological conditions in Filled sites of the Twentymile Creek watershed are not representative of the range of conditions in the entire MTM/VF study area. As a result, comparisons among EIS classes during autumn 2000 and winter 2001 should not be considered typical for the entire MTM/VF study area.

Statistical differences between the Unmined and Filled EIS classes corresponded to ecological differences between classes based on mean WVSCI scores. Unmined sites scored in the Very Good condition category in all seasons except autumn 1999 when the condition was scored as Good. The conditions at Filled sites ranged from Fair to Good (Figure 5-1). However, Filled sites that scored Good on average only represented conditions in the Twentymile Creek watershed in two seasons (i.e., autumn 2000 and winter 2001), and these sites are not representative of the entire MTM/VF study area. On average Filled sites were in worse ecological condition than were Unmined sites.

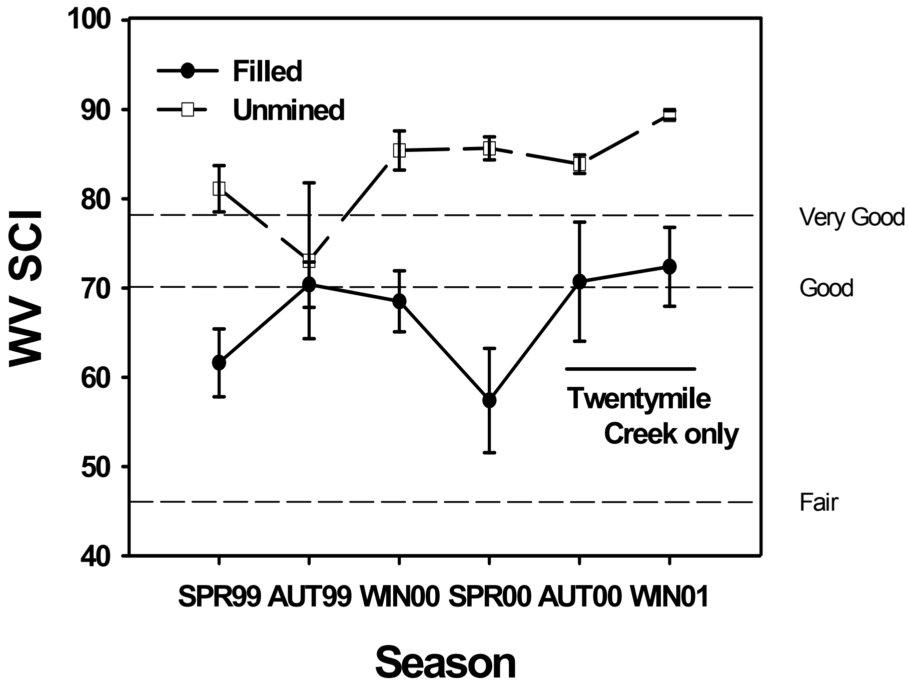


Figure 5-1. Mean WVSCI scores in the Unmined and Filled EIS classes versus sampling season. Error bars are 1 SE. Data for autumn 2000 and winter 2001 only used private organization data for the Twentymile Creek Watershed. The condition categories are based on Green et al. (2000 Draft).

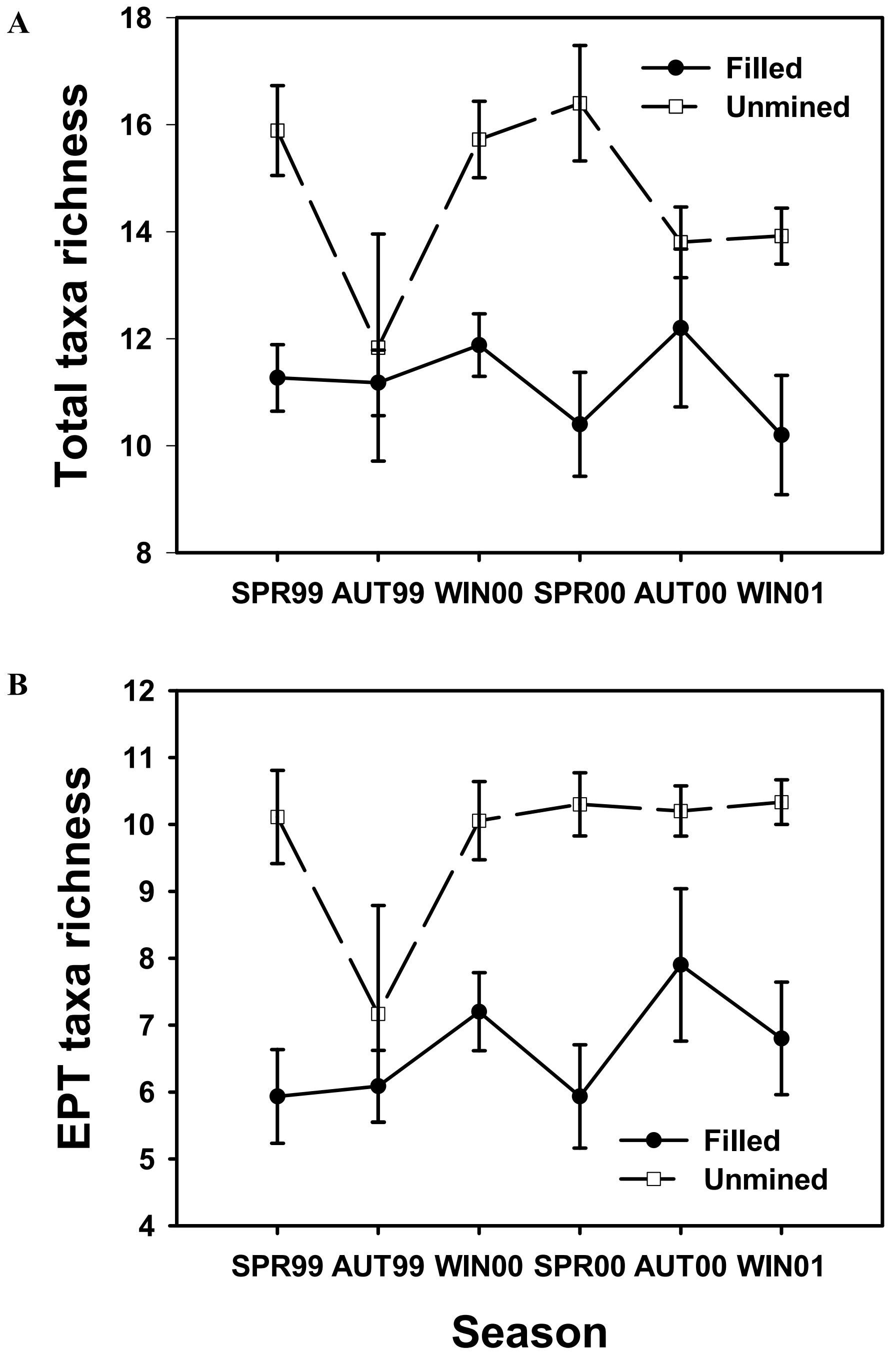


Figure 5-2. (A) Mean Total Taxa richness in the Unmined and Filled EIS classes versus sampling season. (B) Mean EPT Taxa richness in the Unmined and Filled EIS classes versus sampling season. Error bars are 1 SE. Data for autumn 2000 and winter 2001 only used private organization data for the Twentymile Creek Watershed.

The consistently higher WVSCI scores and the Total Taxa in the Unmined sites relative to Filled sites across six seasons showed that Filled sites have lower biotic integrity than those sites without VFs. Furthermore, reduced taxa richness in Filled sites is primarily the result of fewer pollution-sensitive EPT taxa. The lack of significant differences between these two EIS classes in autumn 1999 appears to be due to the effects of greatly reduced flow in sites draining unmined sites during a severe drought. Continued sampling in Unmined and Filled sites would improve the understanding of whether MTM/VF activities are associated with seasonal variation in benthic macroinvertebrate metrics and base-flow hydrology.

Examination of the Additive sites from the mainstem of Twentymile Creek indicated that impacts to the benthic macroinvertebrate communities increased across seasons and upstream to downstream of Twentymile Creek. In the first sampling season one metric, Total Taxa, was negatively correlated with distance along the mainstem. The number of metrics showing a relationship with cumulative river mile increased across seasons, with four of the six metrics having significant correlations in the final sampling season, Winter 2001. Also in Winter of 2001, a regression of the WVSCI versus cumulative river kilometer estimates a decrease of approximately one point in the WVSCI for each river kilometer. Season and cumulative river kilometer in this dataset may be surrogates for increased mining activity in the watershed.

6. LITERATURE CITED

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. *Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates and fish*, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Critchley, M. 2001. Cumulative Hydrologic Impact Assessment of East Fork of Twelvepole Watershed. Department of Environmental Protection, Mining and Reclamation.

Green, J., M. Passemore, and H. Childers. 2000. A Survey of the Condition of Streams in the Primary Region of Mountaintop Mining/Valley Fill Coal Mining (Draft). U.S. Environmental Protection Agency, Region 3 - ESD, Aquatic Biology Group. Wheeling, WV.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monogr. 54:187-211

ITIS, the Integrated Taxonomic Information System, www.itis.usda.gov.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters. A method and its rationale. Illinois Natural History Survey, Special Publication 5. 28p.

Kaufmann, P.R. 1998. *Stream discharge*. Pages 67-76 In J.M. Lazorchak, D.J. Klemm, and D.V. Peck (eds.). Environmental Monitoring and Assessment Program - Surface Waters: Field operations and methods for measuring the ecological condition of wadeable streams. EPA/620/R-94/004F. U.S. Environmental Protection Agency, Washington, D.C.

Kaufmann, P.R., and E.G. Robison. 1998. *Physical Habitat Characterization*. Pages 77-118 In J.M. Lazorchak, D.J. Klemm, and D.V. Peck (eds.). Environmental Monitoring and Assessment Program - Surface Waters: Field operations and methods for measuring the ecological condition of wadeable streams. EPA/620/R-94/004F. U.S. Environmental Protection Agency, Washington, D.C.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger, and D.V. Peck. 1999. Quantifying Physical Habitat in Wadeable Streams. EPA/620/R-99/003. U.S. Environmental Protection Agency, Washington, D.C.

Lazorchak, J.M., D.J. Klemm, and D.V. Peck (eds.). 1998. Environmental Monitoring and Assessment Program - Surface Waters: Field operations and methods for measuring the ecological condition of wadeable streams. EPA/620/R-94/004F. U.S. Environmental Protection Agency, Washington, D.C.

McCormick, F.H., and R.M. Hughes. 1998. *Aquatic Vertebrates*. Pages 161-1181 In J.M. Lazorchak, D.J. Klemm, and D.V. Peck (eds.). Environmental Monitoring and Assessment Program - Surface Waters: Field operations and methods for measuring the ecological condition of wadeable streams. EPA/620/R-94/004F. U.S. Environmental Protection Agency, Washington, D.C.

McCormick, F.H., R.M. Hughes, P.R. Kaufmann, D.P. Peck, J.L. Stoddard, and A.T. Herlihy. 2001. Development of an index of biotic integrity for the Mid-Atlantic highlands region. Trans. Am. Fish. Soc. 130:857-877

Merritt, R.W., and K.W. Cummins. 1996. *An introduction to the aquatic insects of North America*. Third Edition. Kendall/Hunt Publishing Company.

Pecharsky, B.L., P.R. Fraissinet, M.A. Penton, and D.J. Conklin, Jr. 1990. *Freshwater macroinvertebrates of northeastern North America.* Comstock Publishing Associates, Cornell University Press, Ithaca, New York. 442 pp.

Pennak, R.W. 1989. *Fresh-water invertebrates of the United States. Protozoa to Mollusca*. Third Edition. John Wiley & Sons, Inc.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. *Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish.* U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, D.C. EPA 440-4-89-001.

REI Consultants, Inc. 2001. *Comprehensive Quality Assurance Plan*. REI Consultants, Inc., Beaver, West Virginia.

Robins, C. R., R. M. Bailey, C.E. Bond, J.R. Brooker, E.A. Lachner, R.N. Lea, and W.B. Scott (Committee on Names of Fishes). 1991. *Common and Scientific Names of Fishes from the United States and Canada.* 5th edition. American Fisheries Society Special Publication 20.

Stauffer, J.R., Jr., J.M. Boltz, and L.R. White. 1995. *The fishes of West Virginia*. Academy of Natural Sciences of Philadelphia.

Stauffer, J.R., and C.P. Ferreri. 2000. A survey of fish populations in the primary region of mountain top removal/valley fill coal mining. Draft. Pennsylvania State University, School of Forest Resources.

Stewart, K. W., and B. P. Stark. 1988. *Nymphs of North American Stoneflies*. Entomological Society of America. Thomas Say Foundation 12.

Stewart, K.W., and B.P. Stark. 1993. *Nymphs of North American stonefly Genera (Plecoptera)*. University of North Texas Press, Denton.

Tetra Tech, Inc. 1999. Ecological Data Application System (EDAS). A User's Manual. Prepared by Tetra Tech, Inc., Owings Mills, MD.

Trautman, M.B. 1981. *The fishes of Ohio*. Revised edition. Ohio State University Press, Columbus.

U. S. Environmental Protection Agency. 1983. In Methods for Chemical Analysis of Water and Wastes. EPA-600/ 4-79-020. U.S. Environmental Protection Agency., Cincinnati, Ohio.

U.S. Environmental Protection Agency. 1990. Biological criteria: national program guidance for surface waters. Office of the Assistant Administrator for Water. Washington, D.C. EPA/440/5-90/004.

U.S. Environmental Protection Agency. 1996. Summary of State Biological Assessment Programs for Streams and Wadeable Rivers. EPA 230-R-96-007. Office of Water, Washington, D.C.

U.S. Environmental Protection Agency Region 3. 1999. Environmental Impact Statement - Work Assignment. <http://www.epa.gov/region3/mtntop/pdf/workasst.pdf>.

U.S. Environmental Protection Agency. 2000. Interim chemical/biological monitoring protocol for coal mining permit applications. U.S. Environmental Protection Agency, Region 3. http://www.epa.gov/region03/mtntop/pdf/interim_monitorprotocol.pdf

U.S. Environmental Protection Agency Region 3. 2000. Mountaintop Mining-Valley Fill EIS; OSM, EPA, COE, FWS with WVDEP: Goals for the EIS and Questions to be Addressed. June 8, 2000. <http://www.epa.gov/region3/mtntop/pdf/questions.pdf>.

U.S. Environmental Protection Agency Region 3. 2001. US EPA Region 3 Mountaintop Mining. September 7, 2001. <http://www.epa.gov/region3/mtntop/index.htm>.

U.S. Environmental Protection Agency Region 3. 2001. A Survey of the Water Quality of Streams in the Primary Region of Mountain Top Removal / Valley Fill Coal Mining (Draft). Environmental Services Division.

U.S. Environmental Protection Agency. 2002. A survey of the water quality of streams in the primary region of MTM/VF coal mining, WV: October 1999 - January 2001. Draft. U.S. EPA Region 3, Wheeling, WV.

Van Deventer, J.S., and W.S. Platts. 1983. Sampling and estimating fish populations from streams (Statistical methods). Transactions of the North American Wildlife and Natural Resources Conference. 1983. (48th). pp. 349-354.

Walsh, S.J., and M.R. Meador. 1998. Guidelines for quality assurance and quality control for fish taxonomic data collected as part of the National Water Quality Assessment Program. U.S. Geological Survey, Water-Resources Investigations Report 98-4239. Raleigh, NC.

Westfall, M.J., Jr., and M.L. May. 1996. *Damselflies of North America*. Scientific Publishers, Inc.

Wiggins, G.B. 1995. *Larvae of the North American caddisfly Genera (Trichoptera)*. Second edition. University of Toronto Press.

Woods, A.J., J.M. Omernik, D.D. Brown. 1999. Level III and IV Ecoregions of Delaware, Maryland, Pennsylvania, Virginia, and West Virginia. Dynamac Corporation and USEPA. Corvallis, OR.

Wiley, J.B., R.D. Evaldi, J.H. Eychaner, and D.B. Chambers. 2001. Reconnaissance of stream geomorphology, low streamflow, and stream temperature in the mountaintop coal mining region, southern West Virginia, 1999-2000. U.S. Geological Survey, Water Resource Investivations, Report 01-4092. 34 pp.

Yoder, C.O., and E.T. and Rankin 1995. Biological criteria program development and implementation in Ohio. pp 109-144 In W. Davis and T. Simon, eds. Biological Criteria.

Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. Biometrics. 12:163-169

APPENDIX A

SUMMARY TABLES OF PROTOCOLS AND PROCEDURES USED BY THE FOUR ORGANIZATIONS TO COLLECT DATA FOR THE MTM/VF STUDY

Table A-1. Habitat assessment procedures used by the four organizations participating in the MTM/VF Study.

Habitat Assessment Procedures				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Site Selection Criteria	The watershed to be assessed began at least one receiving stream downstream of the mining operation and extended to the headwaters. Monitoring stations were positioned downstream in a similar watershed representative of the future impact scenario. Where possible, semi-annual samples were taken where baseline data were collected. Following Phase II, but prior to final release, samples to be taken where mining phase data were collected. See benthic macroinvertebrate procedures for further details.	No information on habitat data collection given.	Based on agreement reached between the client and regulatory agencies. Sites were selected to provide quantitative, site specific identification and characterization of sources of point and non-point chemical contamination.	No information on habitat data collection given.
Methods Used	Habitat assessment made according to Barbour et al. (1999). Riparian habitat and substrate described using Kaufmann and Robison (1998). Habitat assessment is made as a part of the benthic macroinvertebrate survey.	No information on habitat data collection given.	Habitat assessments performed at the same reach from which biological sampling was conducted. Used the protocols in Kaufmann and Robison (1998) or Barbour et al. (1999).	No information on habitat data collection given.
Procedures	A habitat assessment made according to Barbour et al. (1999) and the riparian habitat and substrate described using Kaufmann and Robison (1998).	No information on habitat data collection given.	A single habitat assessment form which incorporated the features of the sampling reach and of the catchment area was completed. Habitat evaluations were made first on instream habitat, followed by channel morphology, bank structural features and riparian vegetation.	No information on habitat data collection given.

(Continued)

Table A-1. Continued.

Habitat Assessment Procedures (Continued)				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Habitat QA/QC	A habitat assessment made according to Barbour et al. (1999) and the riparian habitat and substrate described using Kaufmann and Robison (1998).	No information on habitat data collection given.	Accepted QA/QC practices were employed during habitat assessment. The habitat evaluations were conducted by a trained field biologist immediately following the biological and water quality sampling. The completed habitat assessment form was reviewed by a second field biologist before leaving the sampling reach. The biologists discussed the assessment. Photographs of the sampling reaches were collected and used as a basis for checks of the assessments. The habitat data were entered into a database, then they were checked against the field sheets.	No information on habitat data collection given.

Table A-2. Parameters and condition categories used in the U.S. EPA's RBP for habitat.

RBP Habitat Parameter	Condition Category			
	Optimal	**Sub-optimal**	**Marginal**	**Poor**
1. Epifaunal Substrate/ Available Cover **(high and low gradient)**	Greater than 70% (50% for low gradient streams) of substrate favorable for epifaunal colonization and fish cover; mix of snags, submerged logs, undercut banks, cobble or other stable habitat and at stage to allow full colonization potential (i.e., logs/ snags that are <u>not</u> new fall and <u>not</u> transient).	40-70% (30-50% for low gradient streams) mix of stable habitat; well-suited for full colonization potential; adequate habitat for maintenance of populations; presence of additional substrate in the form of new fall, but not yet prepared for colonization (may rate at high end of scale).	20 - 40% (10-30% for low gradient streams) mix of stable habitat; habitat availability less than desirable; substrate frequently disturbed or removed.	Less than 20% (10% for low gradient streams) stable habitat; lack of habitat is obvious; substrate unstable or lacking.
SCORE	**20 19 18 17 16**	**15 14 13 12 11**	**10 9 8 7 6**	**5 4 3 2 1 0**
2. Embeddedness **(high gradient)**	Gravel, cobble, and boulder particles are 0-25% surrounded by fine sediment. Layering of cobble provides diversity of niche space.	Gravel, cobble, and boulder particles are 25-50% surrounded by fine sediment.	Gravel, cobble, and boulder particles are 50-75% surrounded by fine sediment.	Gravel, cobble, and boulder particles are more than 75% surrounded by fine sediment.
SCORE	**20 19 18 17 16**	**15 14 13 12 11**	**10 9 8 7 6**	**5 4 3 2 1 0**
3. Velocity/Depth Regimes **(high gradient)**	All four velocity/depth regimes present (slow-deep, slow- shallow, fast-deep, fast-shallow). (Slow is <0.3 m/s, deep is >0.5 m).	Only 3 of the 4 regimes present (if fast-shallow is missing, score lower than if missing other regimes).	Only 2 of the 4 habitat regimes present (if fast-shallow or slow-shallow are missing, score low).	Dominated by 1 velocity/depth regime (usually slow-deep).
SCORE	**20 19 18 17 16**	**15 14 13 12 11**	**10 9 8 7 6**	**5 4 3 2 1 0**
4. Sediment Deposition **(high and low gradient)**	Little or no enlargement of islands or point bars and less than 5% (<20% for low-gradient streams) of the bottom affected by sediment deposition.	Some new increase in bar formation, mostly from gravel, sand or fine sediment; 5-30% (20-50% for low-gradient) of the bottom affected; slight deposition in pools.	Moderate deposition f new gravel, sand or fine sediment on old and new bars; 30-50% 50-80% for low-gradient) of the bottom affected; sediment deposits at obstructions, constrictions, and bends; moderate deposition of pools prevalent.	Heavy deposits of fine material, increased bar development; more than 50% (80% for low-gradient) of the bottom changing frequently; pools almost absent due to substantial sediment deposition.
SCORE	**20 19 18 17 16**	**15 14 13 12 11**	**10 9 8 7 6**	**5 4 3 2 1 0**
5. Channel Flow Status **(high and low gradient)**	Water reaches base of both lower banks, and minimal amount of channel substrate is exposed.	Water fills >75% of the available channel; or <25% of channel substrate is exposed.	Water fills 25-75% of the available channel, and/or riffle substrates are mostly exposed.	Very little water in channel and mostly present as standing pools.
SCORE	**20 19 18 17 16**	**15 14 13 12 11**	**10 9 8 7 6**	**5 4 3 2 1 0**

(Continued)

Table A-2 (Continued).

6. Channel Alteration **(high and low gradient)**	Channelization or dredging absent or minimal; stream with normal pattern.	Some channelization present, usually in areas of bridge abutments; evidence of past channelization (i.e., dredging, greater than past 20 yr) may be present, but recent channelization is not present.	Channelization may be extensive; embankments or shoring structures present on both banks; and 40 to 80% of stream reach channelized and disrupted.	Banks shored with gabion or cement; over 80% of the stream reach channelized and disrupted. In-stream habitat greatly altered or removed entirely.
SCORE	**20 19 18 17 16**	**15 14 13 12 11**	**10 9 8 7 6**	**5 4 3 2 1 0**
7. Frequency of Riffles (or bends) **(high gradient)**	Occurrence of riffles relatively frequent; ratio of distance between riffles divided by width of the stream <7:1 (generally 5 to 7); variety of habitat is key. In streams where riffles are continuos, placement of boulders or other large, natural obstruction is important.	Occurrence of riffles infrequent; distance between riffles divided by the width of the stream is between 7 and 15.	Occasional riffle or bend; bottom contours provide some habitat; distance between riffles divided by the width of the stream is between 15 and 25.	Generally all flat water or shallow riffles; poor habitat; distance between riffles divided by the width of the stream is a ratio of >25.
SCORE	**20 19 18 17 16**	**15 14 13 12 11**	**10 9 8 7 6**	**5 4 3 2 1 0**
8. Bank Stability (score each bank) **(high and low gradient)**	Banks stable: evidence of erosion or bank failure absent or minimal; little potential for future problems. <5% of bank affected.	Moderately stable; infrequent, small areas of erosion mostly healed over. 5-30% of bank in reach has areas of erosion.	Moderately unstable; 30-60% of bank in reach has areas of erosion; high erosion potential during floods.	Unstable; many eroded areas; "raw" areas frequent along straight sections and bends; obvious bank sloughing; 60-100% of bank has erosional scars.
SCORE_____ LB	**Left Bank 10 9**	**8 7 6**	**5 4 3**	**2 1 0**
SCORE_____ RB	**Right Bank 10 9**	**8 7 6**	**5 4 3**	**2 1 0**
9. Bank Vegetative Protection (score each bank) **(high and low gradient)**	More than 90% of the stream bank surfaces and immediate riparian zone covered by native vegetation, including trees, understory shrubs, or nonwoody macrophytes; vegetative disruption through grazing or mowing minimal or not evident; almost all plants allowed to grow naturally.	70-90% of the stream bank surfaces covered by native vegetation, but one class of plants is not well represented; disruption evident but not affecting full plant growth potential to any great extent; more than one-half of the potential plant stubble height remaining.	50-70% of the stream bank surfaces covered by vegetation; disruption obvious; patches of bare soil or closely cropped vegetation common; less than one half of the potential plant stubble height remaining.	Less than 50% of the stream bank surfaces covered by vegetation; disruption of stream bank vegetation is very high; vegetation has been removed to 5 centimeters or less in average stubble height.
SCORE_____ LB	**Left Bank 10 9**	**8 7 6**	**5 4 3**	**2 1 0**
SCORE_____ RB	**Right Bank 10 9**	**8 7 6**	**5 4 3**	**2 1 0**

(Continued)

Table A-2 (Continued).

10. Riparian Vegetation Zone Width (score each bank riparian zone) **(high and low gradient)**	Width of riparian zone >18 meters; human activities (i.e., parking lots, roadbeds, clear-cuts, lawns, or crops) have not impacted zone.	Width of riparian zone 12-18 meters; human activities have impacted zone only minimally.	Width of riparian zone 6-12 meters; human activities have impacted zone a great deal.	Width of riparian zone <6 meters; little or no riparian vegetation due to human activities.
SCORE_____ LB	**Left Bank 10 9**	**8 7 6**	**5 4 3**	**2 1 0**
SCORE_____ RB	**Right Bank 10 9**	**8 7 6**	**5 4 3**	**2 1 0**

Table A-3. Substrate size classes and class scores.

Class	Size	Class Score	Description
Bedrock	> 4000 mm	6	Bigger than a car
Boulder	250 to 4000 mm	5	Basketball to car
Cobble	64 to 250 mm	4	Tennis ball to basketball
Coarse Gravel	16 to 64 mm	3.5	Marble to tennis ball
Fine Gravel	2 to 16 mm	2.5	Ladybug to marble
Sand	0.06 to 2 mm	2	Gritty between fingers
Fines	< 0.06 mm	1	Smooth, not gritty

Table A-4. Water quality assessment procedures used by the four organizations participating in the MTM/VF Study.

Water Quality Procedures				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Site Selection Criteria	The watershed to be assessed began at least one receiving stream downstream of the mining operation and extended to the headwaters. Monitoring stations were positioned downstream in a similar watershed representative of the future impact scenario. Where possible, semi-annual samples were taken where baseline data were collected. Following Phase II, but prior to final release, samples to be taken where mining phase data were collected. See benthic macroinvertebrate procedures for further details.	No information on water quality assessment given.	Based on agreement reached between the client and regulatory agencies. Sites were selected to provide quantitative, site specific identification and characterization of sources of point and non-point chemical contamination.	Not specified in Comprehensive QA Plan.
Methods Used to Make Water Quality Measurements in the Field	Stream flow was measured. Temperature, pH, DO, and conductivity were also measured.	No information on water quality assessment given.	Stream flow was measured at or near the sampling point using techniques in Kaufmann (1998). The data were recorded on a field form. Temperature, pH, DO and conductivity measurements were made using protocols in U.S. EPA (1983). These parameters were measured *in situ* at all sites and recorded on field sheets. The measurements were made directly upstream of the biological sampling site.	Characteristics (i.e., size, depth and flow) and site location are recorded.

(Continued)

Table A-4. Continued.

Water Quality Procedures (Continued)				
	U.S, EPA Region 3	**BMI**	**POTESTA**	**REIC**
Sample Collection	Samples were collected in accordance with Title 40, Chapter I, Part 136 of the Code of Federal Regulations.	No information on water quality assessment given.	Field personnel collected grab samples at each station in conjunction with and upstream of benthic macroinvertebrate sampling events. Water samples were labeled in the field. Samples were collected in accordance with Title 40, Chapter I, Part 136 of the Code of Federal Regulations.	Grab samples are collected with a transfer device or with the sample container. Transfer devices are constructed of inert materials. Samples are placed in appropriate containers. Samples are labeled in the field.
Preservation	Samples were preserved in accordance with Title 40, Chapter I, Part 136 of the Code of Federal Regulations.	No information on water quality assessment given.	Samples were preserved in the field	Samples are preserved in the field. Samples are placed in temperature controlled coolers (4º C) immediately after sampling
Laboratory Transfer	No guidance on water sample transport given.	No information on water quality assessment given.	Samples were transferred to a state-certified laboratory for analysis. Chain-of-custody forms accompanied samples to the laboratory.	Samples are delivered to the laboratory as soon as possible. A chain-of-custody record accompanies each set of samples.

(Continued)

Table A-4. Continued.

Water Quality Procedures (Continued)				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Parameters Analyzed in the Laboratory	Recommended Parameters: dissolved iron dissolved manganese dissolved aluminum calcium magnesium sodium potassium chloride total suspended solids total dissolved solids alkalinity acidity sulfate dissolved organic carbon hardness nitrate/nitrite total phosphorous	No information on water sample analyses given.	alkalinity acidity total suspended and dissolved solids sulfate nitrate/nitrite total phosphorus chloride sodium potassium calcium magnesium hardness total iron total and dissolved manganese total and dissolved aluminum total antimony total arsenic total beryllium total cadmium total chromium total copper total lead total mercury total nickel total selenium total silver total thallium total zinc coarse particulate organic matter fine particulate organic matter total organic carbon	Not specified for this project in the QA Plan.
General QA/QC	A QA/QC plan should be developed.	No information on water chemistry QA/QC practices given.	Accepted QA/QC practices are employed during sampling and analysis.	QA/QC practices are detailed in REI Consultants, Inc. (2001).

(Continued)

Table A-4. Continued.

Water Quality Procedures (Continued)				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Field QA/QC	A QA/QC plan should be developed.	No information on water chemistry QA/QC practices given.	Temperature, pH, DO and conductivity measurements are made using protocols in U.S. EPA (1983). Dissolved oxygen and pH meters are calibrated daily. Calibrations are checked after unusual readings and adjusted if needed. All probes are thoroughly rinsed with distilled water after all calibrations and between sampling sites.	No information on field measurement QA/QC practices given.
Sample Collection QA/QC	A QA/QC plan should be developed.	No information on sample collection QA/QC practices given.	All containers and lids are new. All containers, preservatives and holding times meet the requirements given in Title 40 (Protection of the Environment), Part 136 (Guidelines Establishing Test Procedures for the Analysis of Pollutants) of the Code of Federal Regulations. Each container is labeled with the site identification, date and preservative. Chain-of custody forms are filled out for each group of samples and accompany the samples to a state-certified laboratory.	No information on sample collection QA/QC practices given.
Laboratory QA/QC	A QA/QC plan should be developed.	No information on water sample analysis laboratory QA/QC practices given.	The laboratory analysis of water chemistry follows Standard Methods and/or EPA approved methods. Any deviations from these methods are noted.	No information on water sample analysis laboratory QA/QC practices given.

Table A-5. Fish assemblage assessment procedures used by the four organizations participating in the MTM/VF Study.

Fish Procedures				
	U.S. EPA Region 3 (PSU)	**BMI**	**POTESTA**	**REIC**
Site Selection Criteria	At least one site was established at the most downstream extent of the impact area. This site was permanently recorded and revisited annually. See benthic macroinvertebrate procedures for further details.	No information on fish data collection given.	Sites were designated in consultation with regulatory agencies.	1) Within vicinity of macroinvertebrate and water quality sampling locations. 2) Reaches contained variety of habitat, cover, water velocities and depths. 3) Representative of the stream. 4) If bracketing a confluence, were as close to the tributary as possible, while allowing a downstream buffer for mixing. 5) If used for comparative purposes, contained similar amounts of fish habitat and cover and frequency of riffles and pools.
Station Preparation	Protocols generally followed those in McCormick and Hughes (1998). The stream reach was 40 times the wetted width of the stream, with a maximum reach of 150 m.	No information on fish data collection given.	Stream reach lengths were at least 40 times the stream width and did not exceed 150m.	A stream reach of 150 m was used. Block nets of ⅛-in mesh were set perpendicular to stream by approaching from the shore. Nets were set tight against the substrate and remained in place throughout the survey.
Electrofishing Procedures	Protocols generally followed those in McCormick and Hughes (1998). Block nets were set at the ends of the reach. Amps, voltage and pulse were set according to the stream's conductivity. The surveys began at the downstream end of the reach and proceeded upstream. Netters retrieved the fish and placed them in buckets. The fish were processed at the end of each transect. The survey proceeded until all transects had been fished.	No information on fish data collection given.	Fish were collected at each site using a backpack electrofishing unit. Collections began at the downstream end of the reach and proceeded upstream for the entire reach. Fish collected during the first pass were placed in a bottle labeled "Collection #1". Two additional passes were made and fish from the second and third pass were placed in bottles labeled "Collection #2" and "Collection #3, respectively. If the number of fish in the latter passes did not decline from the previous pass, additional passes were made.	Surveys were conducted in first-, second- and third-order streams by a backpack electrofishing unit. The output voltage and pulse frequency were controlled by the biologist. The biologist progressed slowly upstream moving the wands across the entire stream width. Technicians positioned on each side of the biologist netted the stunned fish and placed them in buckets containing water. Three passes were conducted at each station.

(Continued)

Table A-5. Continued.

Fish Procedures (Continued)				
	U.S. EPA Region 3 (PSU)	**BMI**	**POTESTA**	**REIC**
Field Measurements	Fish were identified, tallied and examined for external anomalies. The standard length of each fish was measured to the nearest mm and each fish was weighed to the nearest 0.01 g.	No information on fish data collection given.	Fish from each pass were kept separate. Game fish (except small specimens) and rare, threatened or candidate species were counted, measured (total length), weighed and released. These data were recorded on field sheets. The majority of fish captured were preserved in 10% formalin and taken to the laboratory. Each collection was preserved separately.	After each pass, fish were identified, measured to the nearest mm of total length and weighed to the nearest 0.1 gm or 1.0 gm (depending on fish size). Large fish were held in a live well until the completion of the survey, then released to their original reach. Small fish requiring microscopic verification were preserved in 10% formalin and taken to the laboratory.
Specimen Preparation, Identification and Validation	Fish were labeled and preserved in 10% formalin and transported to the PSU Fish Museum where they were deposited for permanent storage in 50% isopropanol. Voucher collections of up to 25 individuals of each taxon collected (except very large individuals of easily identified species) were prepared.	No information on fish data collection given.	Preserved specimens were taken to the laboratory and temporarily stored in 50% isopropanol or 10% ethanol. They were identified and weighed. All preserved fish were placed in permanent storage in a recognized museum collection or offered for use in the federal EIS on MTR/VF mining in West Virginia.	Small fish were identified in the laboratory. All fish were sorted by species and their identities were verified when they were weighed to the nearest 0.1 gm and their total lengths were measured. Identified fish were stored. Unidentified fish were identified and validated by West Virginia DNR personnel.
Fish Data Analysis	Total biomass caught, biomass per m^2 sampled and abundances of each species were calculated.	No information on fish data analysis given.	Fish data sheets were transferred into spreadsheets. Data entered into the spreadsheets were routinely checked against field and laboratory sheets immediately following data entry. Any discrepancies were documented and corrected. Population and community structure were determined at each site. Age classes based on length, frequency analysis and standing crop (kg/ha) were calculated for each species at each pass.	Data were entered into a spreadsheet and confirmed. At each sampling station, total taxa, number and percent of pollution-intolerant fish, number and percent of intermediately pollution- tolerant fish, Number and percent of pollution-tolerant fish, Shannon-Weiner diversity Index, Percent species similarity index were made. For each species at each sampling station, Total abundance, Mean length, Mean weight, Standing stock, and Sensitivity index (U.S. EPA 1999) were calculated.

(Continued)

Table A-5. Continued.

Fish Procedures (Continued)				
	U.S. EPA Region 3 (PSU)	**BMI**	**POTESTA**	**REIC**
Fish Population Estimates	No information on fish population estimates given.	No information on fish data analysis given.	Population estimates of each species at each site were made using the triple pass depletion method of Van Deventer and Platts (1983).	Population estimates for each species and each reach were calculated using the Zippin (1956) depletion method and based on observed relative abundance. Total fish weight by species was extrapolated to calculate an estimated total standing stock.
Fish Identification and Verification QA/QC	The interim protocols stated that a QA/QC plan should be developed.	No information on fish data QA/QC given.	Implemented the QA/QC plan from the U.S. Geological Survey (Walsh and Meador 1998). The plan outlines methods used to ensure accurate identification of fish collected. A voucher collection including one specimen of each taxon collected was made available for verification. Data entered into spreadsheets were routinely checked against field and laboratory sheets.	The QA/QC protocols called for the use of two Fisheries Biologists with the appropriate qualifications: Any species captured whose distribution did not match Stauffer et al. (1995) was recorded and the identification was confirmed by West Virginia DNR personnel. All identifications were confirmed by both Fisheries Biologists. Small fish which required microscopic identification were stored for future reference or identification. A reference collection of all captured taxa was kept. Any species of questionable identification were kept and verified by West Virginia DNR personnel. All retained specimens were permanently labeled.

Table A-6. Macroinvertebrate assemblage assessment procedures used by the four organizations participating in the MTM/VF Study.

Benthic Macroinvertebrate Procedures				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Site Selection Criteria	The watershed to be assessed began at least one receiving stream downstream of the mining operation and extended to the headwaters. Monitoring stations were positioned downstream in a similar watershed representative of the future impact scenario. Where possible, semi-annual samples were taken where baseline data were collected. A minimum of two stations were established for each intermittent and perennial stream where fills were proposed. One station was as close as possible to the toe of the fill and the other was downstream of the sediment pond location. If the sediment pond was more than 0.25 mi from the toe of the fill, a third station was placed between the two. Additional stations were placed in at least the first receiving stream downstream of the mining operation.	BMI located one sampling station as close as possible to the toe of the proposed VF. Another sampling station was located below the proposed sediment pond. If the proposed sediment pond was to be > 0.25 miles below the toe of the fill, an additional station was located between the toe of the fill and the sediment pond. Two sampling stations were located within the next order receiving stream downstream. One of these stations was located above the confluence and one was located below the confluence. In general, an unmined reference station was located at a point that represented the area proposed for mining. In addition, a mined and filled reference station was located at a point that represents a similar level of mining.	Based on an agreement reached between the client and regulatory agencies. Selected to provide quantitative and qualitative characterizations of benthic macroinvertebrate communities.	The sampling station locations contained habitat which was representative of the overall habitat found within stream reach. Stations that were to be used for comparative purposes contained similar habitat characteristics. Stations bracketing a proposed fill tributary were close (approximately 100 m) to the impacted tributary. The general locations were usually pre-determined by the client and the permit writer. When descriptions of predetermined sites were vague, professional judgements were made in an attempt to incorporate the studies' goals. For selecting sampling sites for proposed VFs, site were located at the toe of the valley, below the sediment pond at the mouth of the fill stream, upstream and downstream of the fill stream on the receiving stream and on the next order receiving stream.

(Continued)

Table A-6. Continued.

Benthic Macroinvertebrate Procedures (Continued)				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Sampling Point selection	The sampling point was at the middle of the reach. It was moved upstream or downstream to avoid tributary effects, bridges or fords.	No information given on specific sampling point selection.	No information given on specific sampling point selection.	One of three methods (i.e., completely randomized, stratified-random or stratified) was used to select the sampling points at a site. Generally, the stratified-random method was used in large streams and the stratified method was used in small streams. In small intermittent streams or when there was little water, samples were taken from wherever possible.
Sampler Used	Sampling was conducted according to Barbour et al. (1999). A 0.5-m rectangular kick net was used to composite four ¼-m^2 samples.	In the autumn of 1999 and the spring of 2000, four ¼-m^2 samples collected with a D-frame kick net were composited. In the autumn of 2000, six Surber samples were collected and four ¼-m^2 samples collected with a D-frame kick net were composited. In the spring of 2001, four Surber samples, were collected and four ¼-m^2 samples were collected with a D-frame kick net and composited.	Four ¼-m^2 samples were taken using a D-frame kick net and composited. Surber samplers were used at selected sampling stations.	The sampling devices were dependent on the permit. Three samples were taken using a Surber sampler. These were not composited. Four ¼-m^2 samples were taken using a D-frame kick net. These were composited. The Surber samplers were usually used in riffle areas and the kick net samples were usually taken from deeper run or pool habitats.
Surber Sampler Procedures	Surber samplers were not used.	The frame of the sampler was placed on the stream bottom in the area that was to be sampled. All large rocks and debris that are in the 1.0-ft^2 frame were scrubbed and rinsed into the net and removed from the sampling area. Then, the substrate in the frame was vigorously disturbed for 20 seconds. Each sample was rinsed and placed into a labeled container with two additional labels inside the sample containers.	The Surber sampler was placed with all sides flat on the stream bed. Large cobble and gravel within the frame were brushed. The area within the frame was disturbed to a depth of three in with the handle of the brush. The sample was transferred to a labeled plastic bottle.	The sampler was placed with the cod end downstream. The substrate upstream of the sampler was scrubbed gently with a nylon brush for up to three minutes. Water was kept flowing into sampler while scrubbing. Rocks were checked and any clinging macroinvertebrates were removed and placed in the sampler. The material in the sampler was rinsed and collected into a bottle.

(Continued)

Table A-6. Continued.

Benthic Macroinvertebrate Procedures (Continued)				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Kick Net Procedures	The procedures in Barbour et al. (1999) were modified so that 1 m^2 of substrate was sampled at each site.	The net was held downstream of the 0.25-m^2 area that was to be sampled. All rocks and debris that were in the 0.25-m^2 area were scrubbed and rinsed into the net and removed from the sampling area. Then, the substrate in the 0.25-m^2 area was vigorously disturbed for 20 seconds. This process was repeated four times at each sampling site. The composited sample was rinsed and placed into a labeled container.	The kick net samples were collected using protocols in Barbour et el. (1999). All boulders, cobble and large gravel within 0.25 m^2 upstream of net were brushed into the net. The substrate within 0.25 m^2 upstream of the net was kicked for 20 seconds. Four samples were collected and composited. The sample was transferred to a labeled plastic bottle.	The sampler was placed with the net outstretched and the cod end downstream. The substrate was kicked or scrubbed for up to three minutes. Discharged material was swept into the net. An area of approximately 0.25m^2 was sampled. The procedure was repeated four times.
Additional information collected from sites	The physical/chemical field sheets were completed before sampling and they were reviewed for accuracy after sampling. A map of the sampling reach was drawn. A GPS unit was used to record latitude and longitude. After sampling, the Macroinvertebrate Field Sheet was completed. The percentage of each habitat type in the reach was recorded and the sampling gear used was noted. Comments were made on conditions of the sampling.. Observations of aquatic flora and fauna were documented. Qualitative estimates of macroinvertebrate composition and relative abundance were made. A habitat assessment was made. Riparian habitat was described using Kaufmann and Robison (1998).	Additional information collected was not described.	A field data sheet (from Barbour et al. 1999) was completed and photographic documentation was taken at the time of sampling. Photographs showed an upstream view and a downstream view from the center of the sampling reach.	Additional information collected was not described.

(Continued)

Table A-6. Continued.

Benthic Macroinvertebrate Procedures (Continued)				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Sample Preservation	Samples were preserved in 95% ethanol.	Samples were preserved in 70% ethanol.	Quantitative samples were preserved in 50% isopropanol. Semi-quantitative samples were preserved in either 50% isopropanol or 70% ethanol.	Samples were preserved in the field with formaldehyde (30% by wt.). Approximately 10% of the samples' volume was added.
Logging samples	All samples were dated and recorded in a sample log notebook upon receipt by laboratory personnel. All information from the sample container label was included on the sample log sheet (Barbour et al. 1999).	Samples were logged onto Chain-of-Custody forms. Logs were maintained throughout the identification process.	When samples arrived at the laboratory, they were entered in a log book and tracked through processing and identification.	Sample logging procedure was not described.
Laboratory Procedures	Samples were thoroughly rinsed in a 500 µm-mesh sieve. Large organic material was rinsed, visually inspected, and discarded. Samples that had been preserved in alcohol, were soaked in water for approximately 15 minutes. Samples stored in more than one container were combined. After washing, the sample was spread evenly across a pan marked with grids approximately 6 cm x 6 cm. A random numbers table was used to select four grids. All material from the four grids (⅛ of the total sample) was removed and placed in a shallow white pan. A predetermined, fixed number of organisms were used to determine when sub-sampling was complete.	Samples were rinsed using a #24 sieve (0.0277-in mesh) and then transferred to an enamel tray. Water was added to the tray to a level that covered the sample. All macroinvertebrates in the sample were picked from the debris using forceps and then transferred to a vial that contained 70% ethanol. One of the labels from the sample jar was placed on the organism vial. After identification and processing, the samples were then stored according to the project plan.	Benthic macroinvertebrates were processed using the single habitat protocols in Barbour et al. (1999). The entire samples were processed. Identifications were recorded on standard forms. Ten percent of the samples are re-picked and identifications are randomly reviewed.	Samples were processed individually. They were poured into a 250-µm sieve. Then rinsed with water and transferred to a four-part sub-sampler with a 500-µm screen and distributed evenly on the with water. The first ¼ of the sample was put into petri dishes and the aquatic insects were sorted from the detritus. All macroinvertebrates were placed in a labeled bottle with formalin. If too few individuals were found in the ¼, the second ¼ was picked. Then, either a portion of the picked detritus was re-checked, or a single sorter checked all petri dishes. If organisms were present, the sample was re-picked. After sample sorting was complete, picked and unpicked detritus was stored.

(Continued)

Table A-6. Continued.

Benthic Macroinvertebrate Procedures (Continued)				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Benthic Macro-invertebrate Identification	Organisms were identified to the lowest practical taxon by a qualified taxonomist. Each taxon found in a sample was recorded and enumerated in a bench notebook and then transcribed to the laboratory bench sheet for subsequent reports. Any difficulties encountered during identification were noted on these sheets. Labels with specific taxa names were added to the vials of specimens. The identity and number of organisms were recorded on the bench sheet. Life stages of organisms were also recorded (Barbour et al. 1999).	Using a binocular compound microscope, each organism was identified to the taxa level specified in the project study plan. The numbers of organisms found in each taxa were recorded on bench sheets. Then, the organisms and sample label were returned to the organism vial and preserved with 70% ethanol. For QC purposes, 10% of all samples were re-identified.	Samples were identified by qualified freshwater macroinvertebrate taxonomists to the lowest practical taxon.	Aquatic insects were identified under a microscope to the lowest practical taxonomic level. Unless specified otherwise, Chironomids were identified to the Family level and Annelids were broken into classes. Identified specimens were returned to the sample bottle and preserved in formalin. New or extraordinary taxa were added to reference collections. Random samples are re-identified periodically.
Macro-invertebrate Sample Storage	Samples were stored for at least six months. Specimen vials were placed in jars with a small amount of 70% ethanol and tightly capped. The ethanol level in these jars was examined periodically and replenished as needed. A label was placed on the outside of the jar indicating sample identifier, date, and preservative.	No information on sample storage was provided.	No information on sample storage was provided.	Samples were stored for at least six months.
Database Construction	No information on database construction was provided.	No information on database construction was provided.	The data from the taxonomic identification sheets were transferred into spreadsheets. Data entered into the spreadsheets were routinely checked against field and laboratory sheets.	No information on database construction was provided.
Benthic Macro-invertebrate Data Analysis	Data were used to calculate the WVSCI.	No information on data analysis was provided.	Eight bioassessment metrics were calculated for each sampling station.	Twelve benthic macroinvertebrate metrics were calculated for each of the sampling stations. Abundance data from sub-sampling was extrapolated to equal the entire sample amount.

(Continued)

Table A-6. Continued.

Benthic Macroinvertebrate Procedures (Continued)				
	U.S. EPA Region 3	**BMI**	**POTESTA**	**REIC**
Benthic Macro-invertebrate Metrics Calculated	Data were used to calculate the metrics of the WVSCI.	No information on metrics was provided.	1. Taxa Richness 2. Total Number of Individuals 3. Percent Mayflies 4. Percent Stoneflies 5. Percent caddisflies 6. Total Number of EPT Taxa 7. Percent EPT Taxa 8. Percent Chironomidae	1. Taxa Richness 2. Modified HBI: Summarizes overall pollution tolerance. 3. Ratio of Scrapers to Filtering Collectors 4. Ratio of EPTs to Chironomidae 5. Percent of Mayflies 6. Percent of Dominant Family 7. EPT Index: Total number of distinct taxa within EPT Orders. 8. Ratio of Shredders to Total Number of Individuals 9. Simpson's Diversity Index 10. Shannon-Wiener Diversity Index 11. Shannon-Wiener Evenness 12. West Virginia Stream Condition Index: a six-metric index of ecosystem health.

APPENDIX B

IBI COMPONENT METRIC VALUES

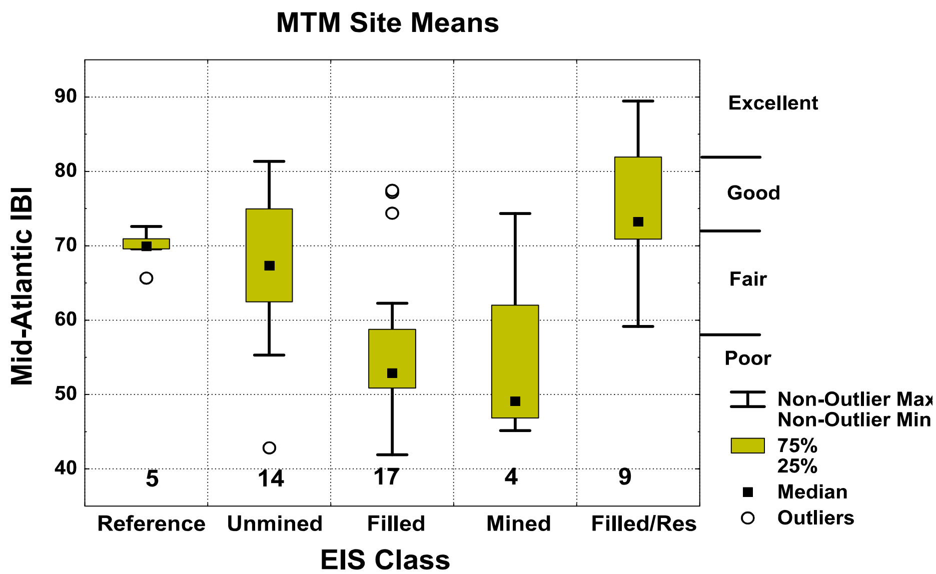


Figure B-1. Box plot of the IBI among EIS classes and regional reference sites. All taxa richness metrics were adjusted to a catchment area of 100 km2.

Table B-1. The ANOVA for IBI scores among EIS classes (Unmined, Filled, Mined, and Filled/Residential).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2335.56	778.52	6.70	0.0009
Error	40	4651.31	116.28		
Corrected Total	43	6986.87			

R-Square	Coefficient of Variance	Root MSE	Index Mean
0.334	17.022	10.783	63.350

Table B-2. Dunnett's test comparing IBI values of EIS classes to the Unmined class, with the alternative hypothesis that IBI < Unmined IBI (one-tailed test).

EIS Class	N	Mean	Standard Deviation	Dunnett's P-Value
Filled	17	56.8	10.6	0.0212
Filled/Residential	9	74.6	10.7	0.9975
Mined	4	54.4	13.4	0.0685
Unmined	14	66.7	10.3	--

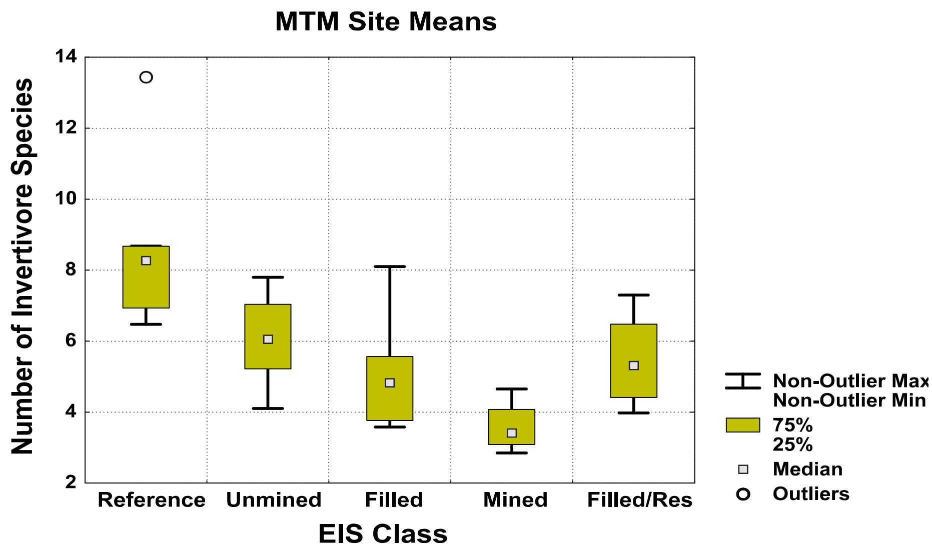


Figure B-2. Box plot of the Number of Benthic Invertivore Species among EIS classes and regional reference sites.

Table B-3. The ANOVA for Number of Benthic Invertivore Species among EIS classes (Unmined, Filled, Mined, and Filled/Residential).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	22.32	7.44	4.91	0.0054
Error	40	60.66	1.51		
Corrected Total	43	82.98			

R-Square	Coefficient of Variance	Root MSE	Index Mean
0.269	23.504	1.231	5.239

Table B-4. Dunnett's test comparing Numbers of Benthic Invertevores to the Unmined class, with the alternative hypothesis that IBI < Unmined IBI (one-tailed test).

EIS Class	N	Mean	Standard Deviation	Dunnett's P-Value
Filled	17	4.8	1.3	0.0182
Filled/Residential	9	5.4	1.2	0.3234
Mined	4	3.6	0.76	0.0017
Unmined	14	6.0	1.2	--

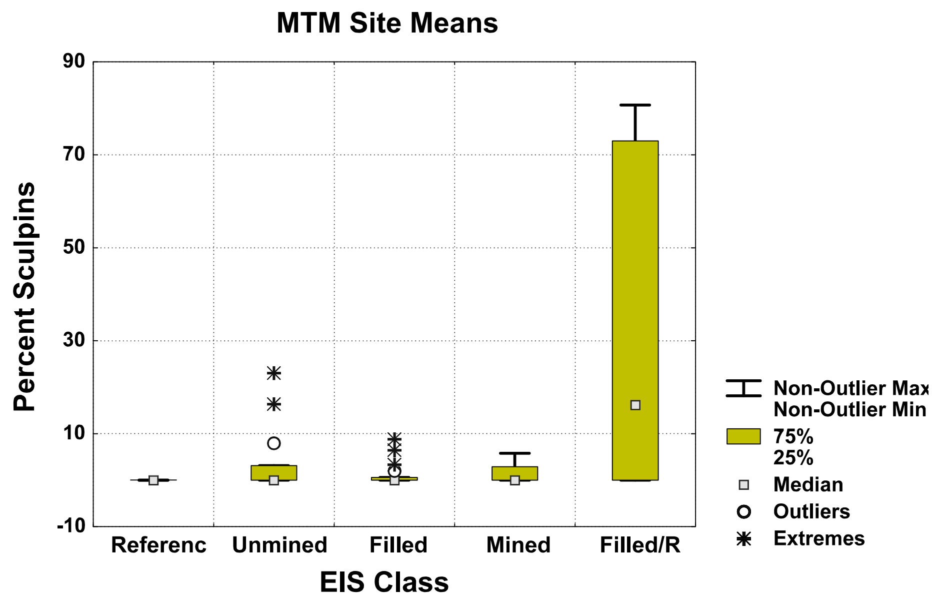


Figure B-3. Box plot of the Percent Cottidae(Sculpins) among EIS classes and regional reference sites.

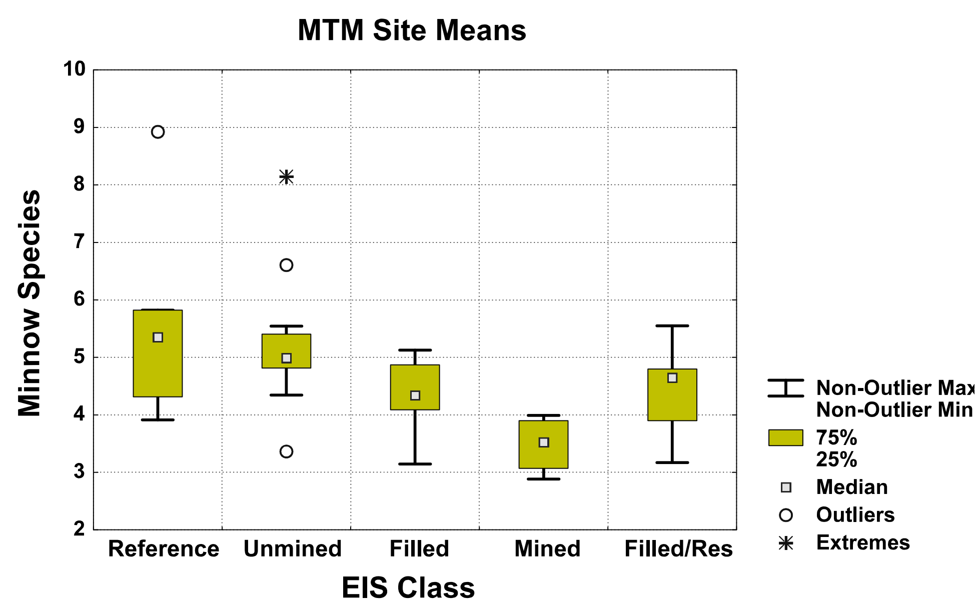


Figure B-4. Box plot of the Number of Native Cyprinidae (Minnow Species) among EIS classes and regional reference sites. This metric was adjusted to a catchment area of 100 km2.

Table B-5. The ANOVA for Number of Native Cyprinidae (Minnow Species) among EIS classes (Unmined, Filled, Mined, and Filled/Residential).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	11.36	3.79	5.79	0.0022
Error	40	26.19	0.65		
Corrected Total	43	37.56			

R-Square	Coefficient of Variance	Root MSE	Index Mean
0.302	17.777	0.809	4.55

Table B-6. Dunnett's test comparing Numbers of Native Cyprinidae (Minnows Species) to the Unmined class, with the alternative hypothesis that IBI < Unmined IBI (one-tailed test).

EIS Class	N	Mean	Standard Deviation	Dunnett's P-Value
Filled	17	4.3	0.58	0.0089
Filled/Residential	9	4.4	0.73	0.0311
Mined	4	3.5	0.51	0.0008
Unmined	14	5.2	1.1	--

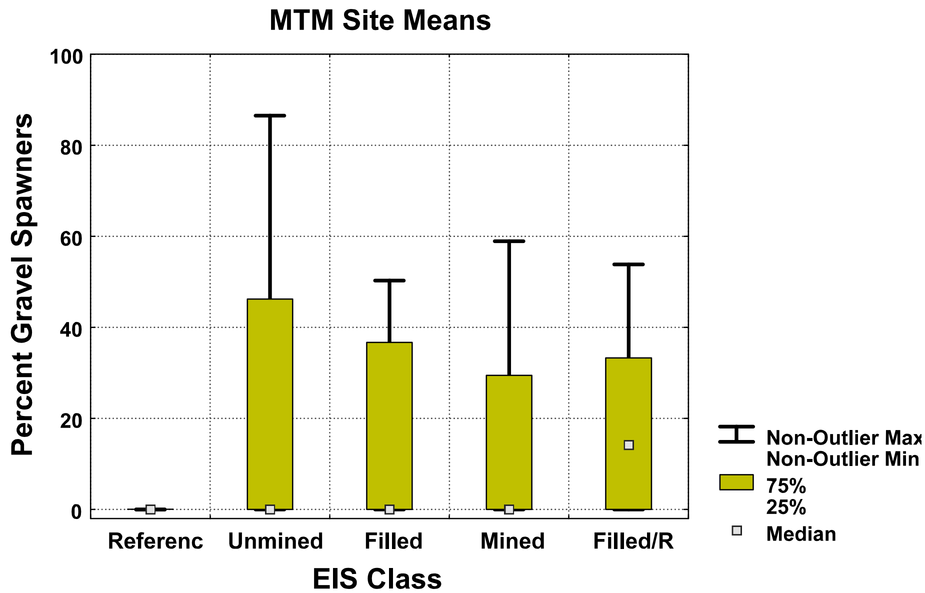


Figure B-5. Box plot of the Percent Gravel Spawners among EIS classes and regional reference sites.

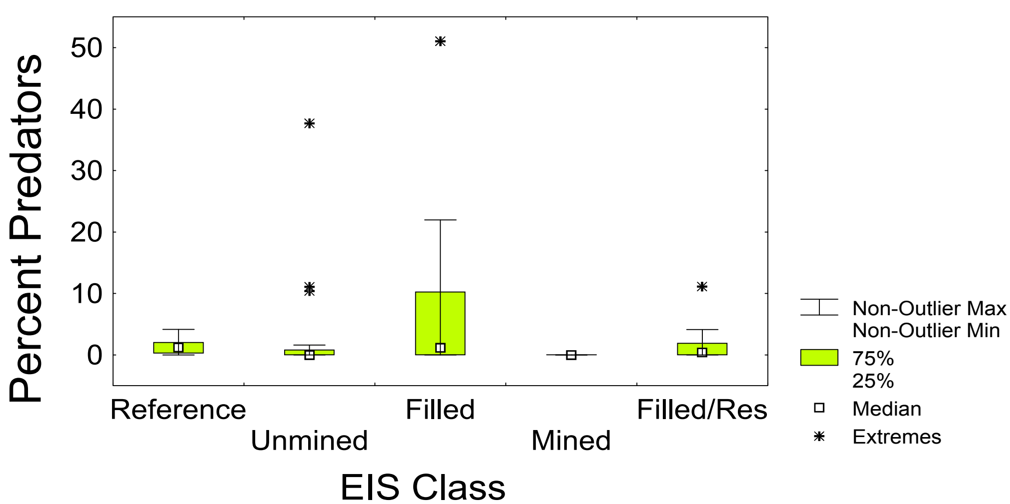


Figure B-6. Box plot of the Percent Piscivore/Invertivores (Predators) among EIS classes and regional reference sites.

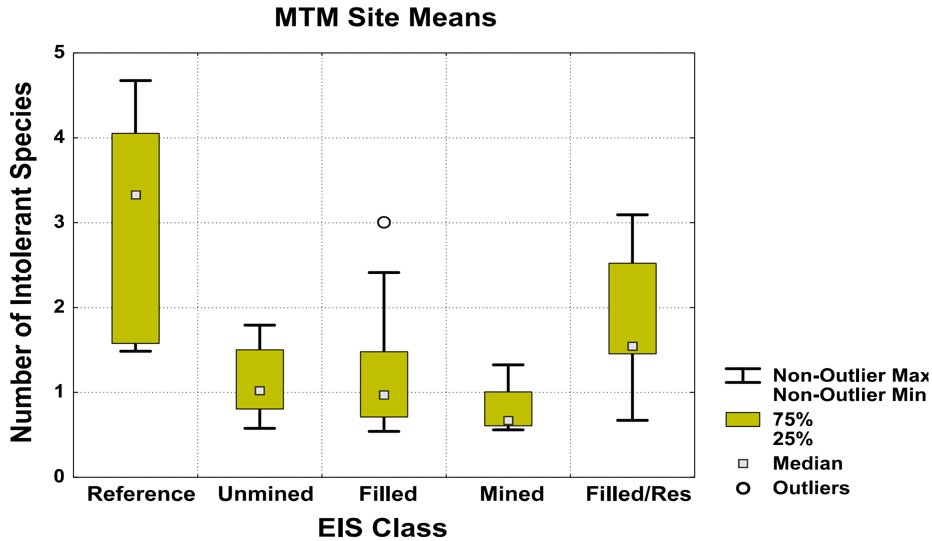


Figure B-7. Box plot of the Number of Intolerant Species among EIS classes and regional reference sites. This metric was adjusted to a catchment area of 100 km2.

Table B-7. The ANOVA for Number of Intolerant Species among EIS classes (Unmined, Filled, Mined, and Filled/Residential).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5.29	1.76	5.96	0.0019
Error	40	11.83	0.29		
Corrected total	43	17.12			

R-Square	Coefficient of Variance	Root MSE	Index Mean
0.308	44.209	0.543	1.23

Table B-8. Dunnett's test comparing Numbers of Intolerants to the Unmined class, with the alternative hypothesis that IBI < Unmined IBI (one-tailed test).

EIS Class	N	Mean	Standard Deviation	Dunnett’s P-Value
Filled	17	1.1	0.49	0.7075
Filled/Residential	9	1.9	0..83	1.0000
Mined	4	0.8	0.35	0.3504
Unmined	14	1.1	0.40	--

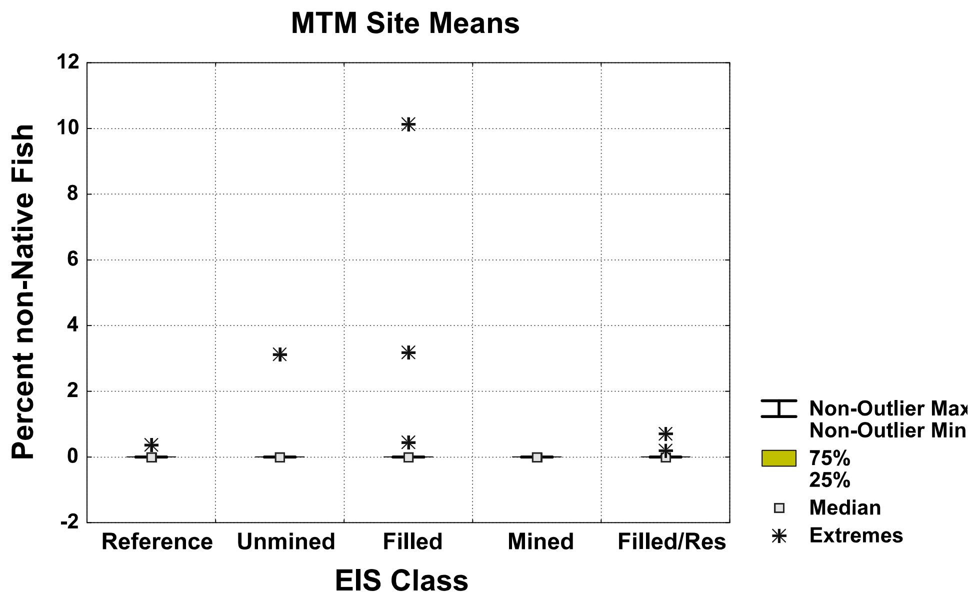


Figure B-8. Box plot of the Percent Exotic (Non-Native Fish) among EIS classes and regional reference sites.

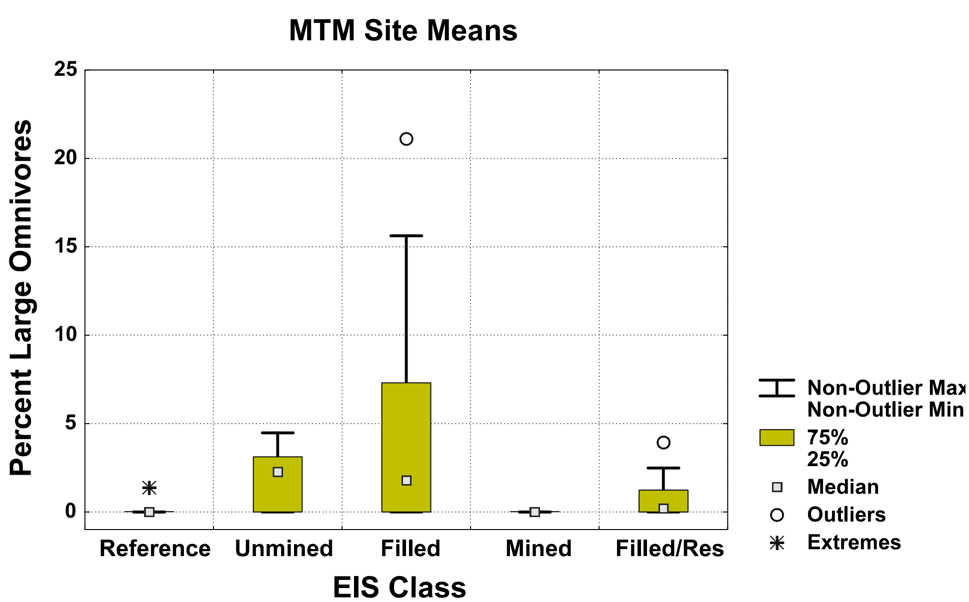


Figure B-9. Box plot of the Percent Macro Omnivores among EIS classes and regional reference sites.

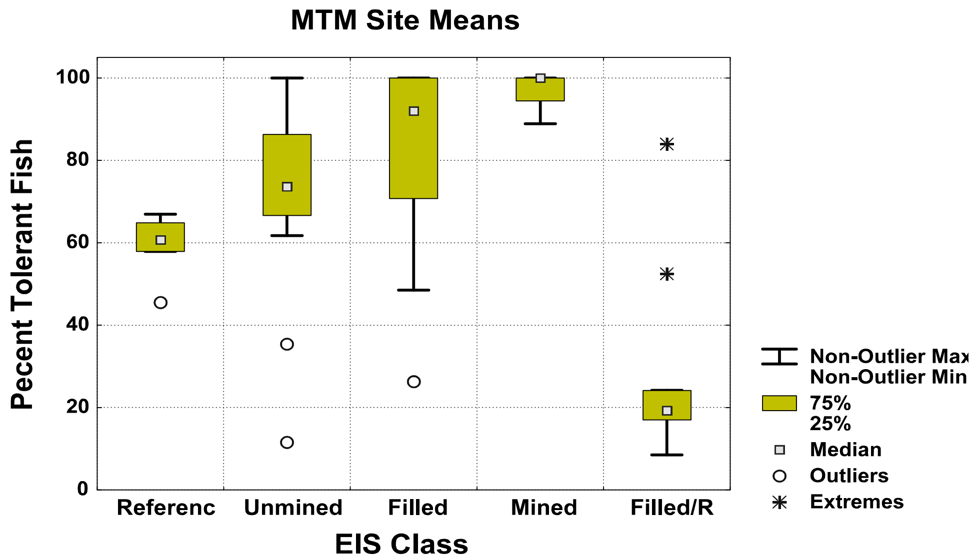


Figure B-10. Box plot of the Percent Tolerant Fish among EIS classes and regional reference sites.

Table B-9. The ANOVA for Number of Tolerant Species among EIS classes (Unmined, Filled, Mined, and Filled/Residential).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	21001.35	7000.45	14.03	<0.0001
Error	40	19956.38	498.91		
Corrected total	43	40957.73			

R-Square	Coefficient of Variance	Root MSE	Index Mean
0.512	32.055	22.336	69.681

Table B-10. Dunnett's test comparing Numbers of Tolerant Species to the Unmined class, with the alternative hypothesis that IBI < Unmined IBI (one-tailed test).

EIS Class	N	Mean	Standard Deviation	Dunnett's P-Value
Filled	17	82.9	21.5	0.2080
Filled/Residential	9	28.9	24.1	1.0000
Mined	4	97.2	5.6	0.0681
Unmined	14	71.8	24.6	--

APPENDIX C

BOX PLOTS OF THE WVSCI AND COMPONENT METRICS

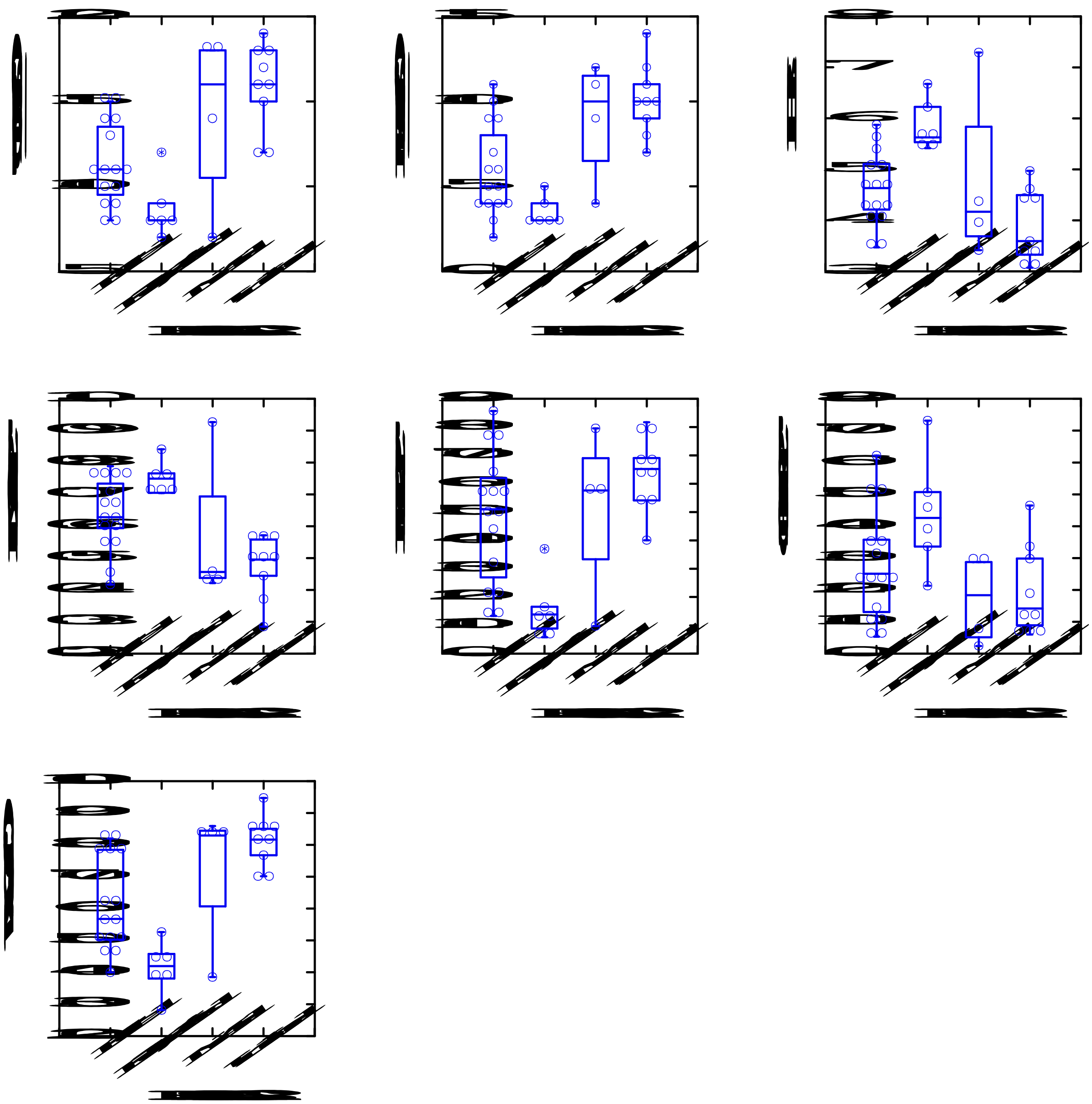


Figure C-1. Box plots of the WVSCI and its component metrics versus the EIS class for the spring 1999 season. Circles represent site scores.

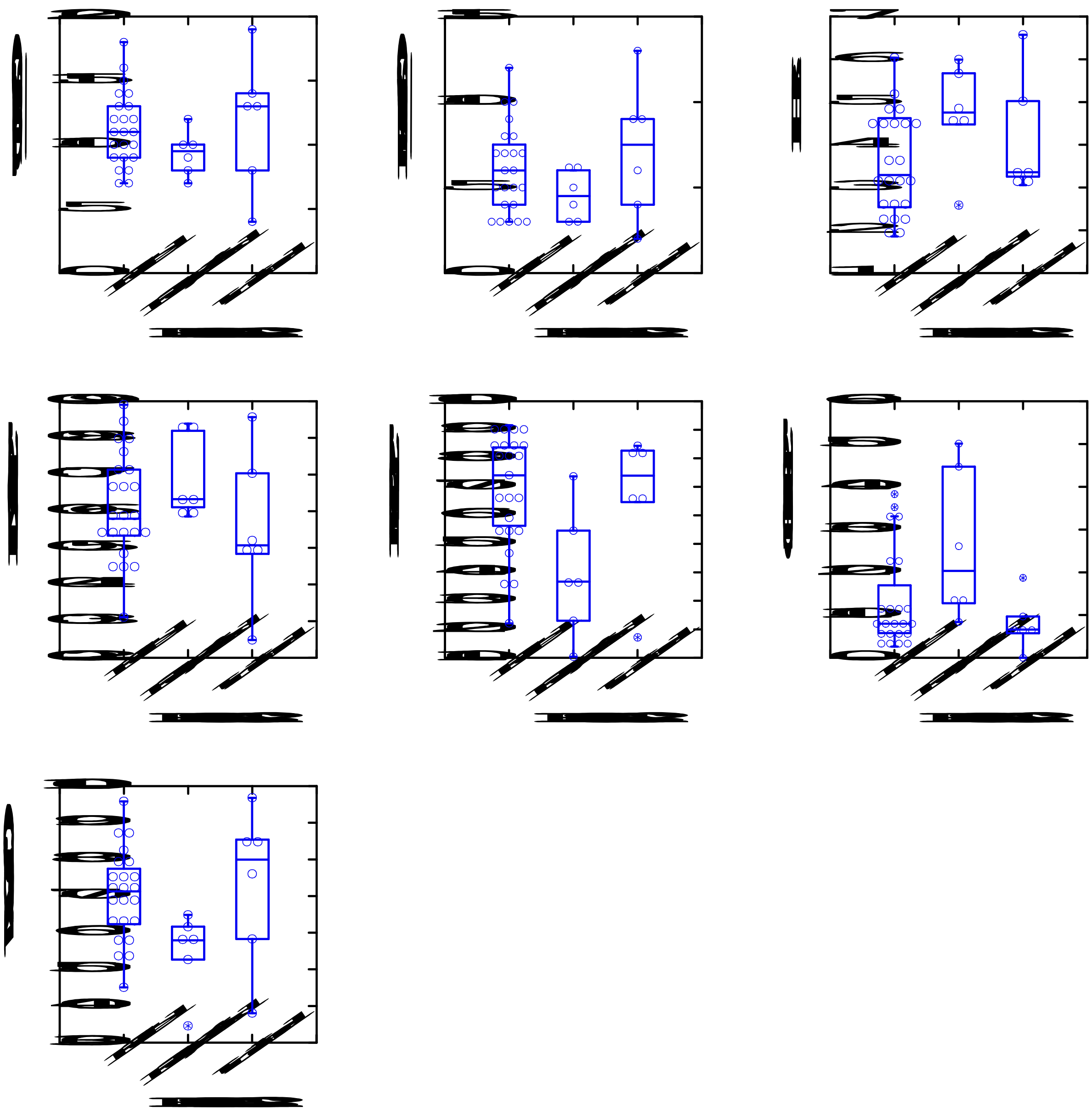


Figure C-2. Box plots of the WVSCI and its component metrics versus the EIS class for the autumn 1999 season. Circles represent site scores.

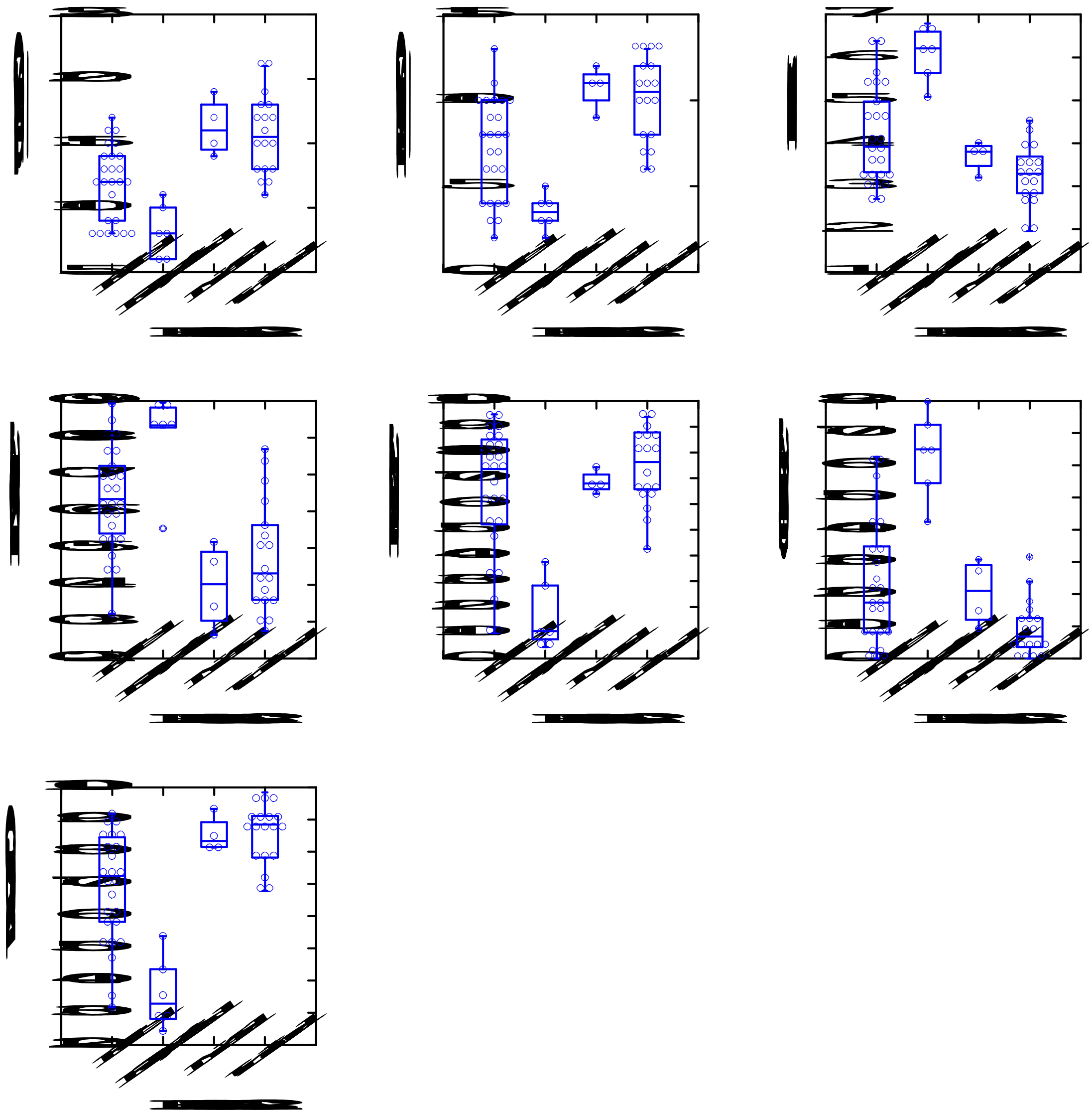


Figure C-3. Box plots of the WVSCI and its component metrics versus the EIS class for the winter 2000 season. Circles represent site scores.

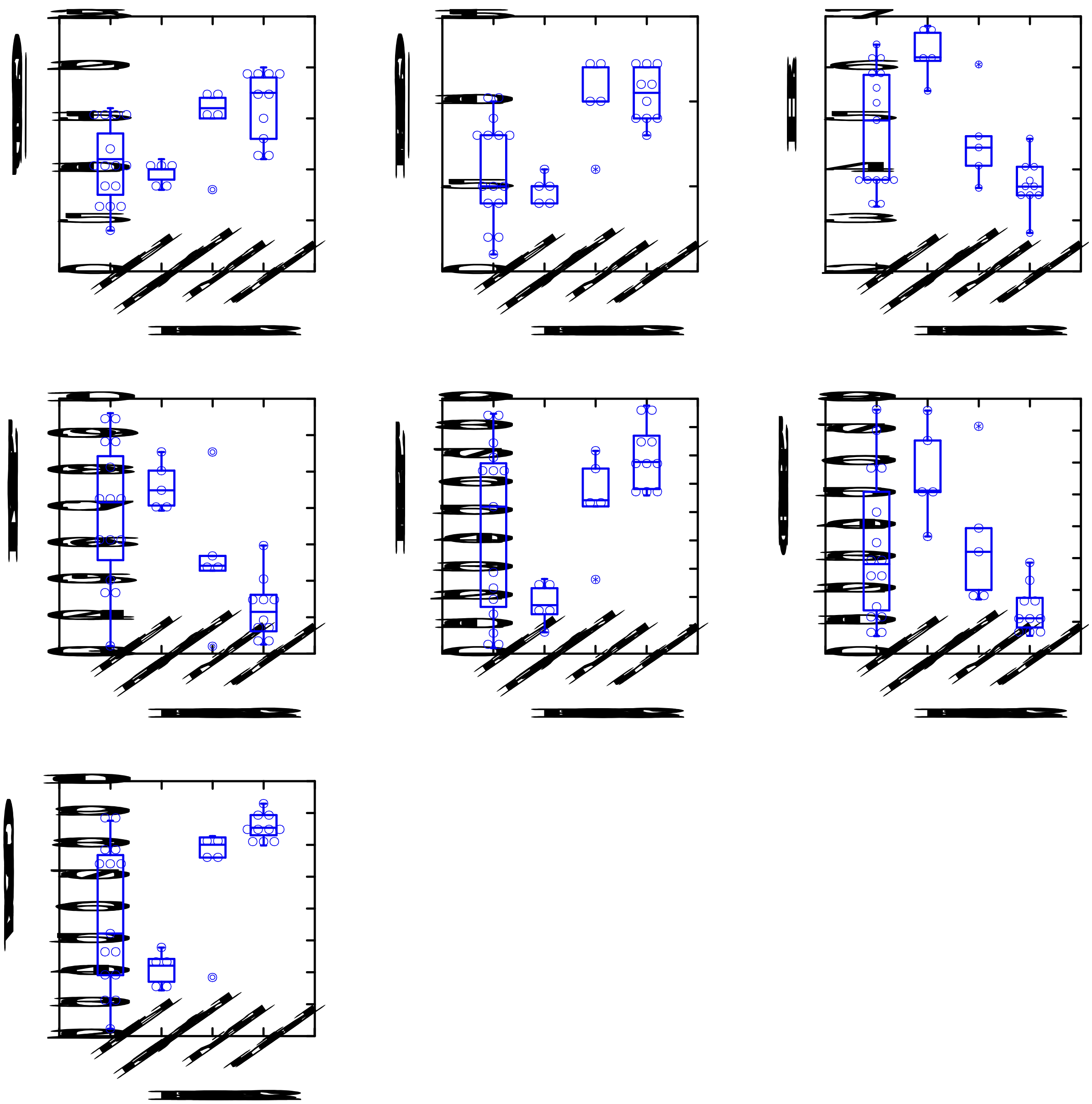


Figure C-4. Box plots of the WVSCI and its component metrics versus the EIS class for the spring 2000 season. Circles represent site scores.

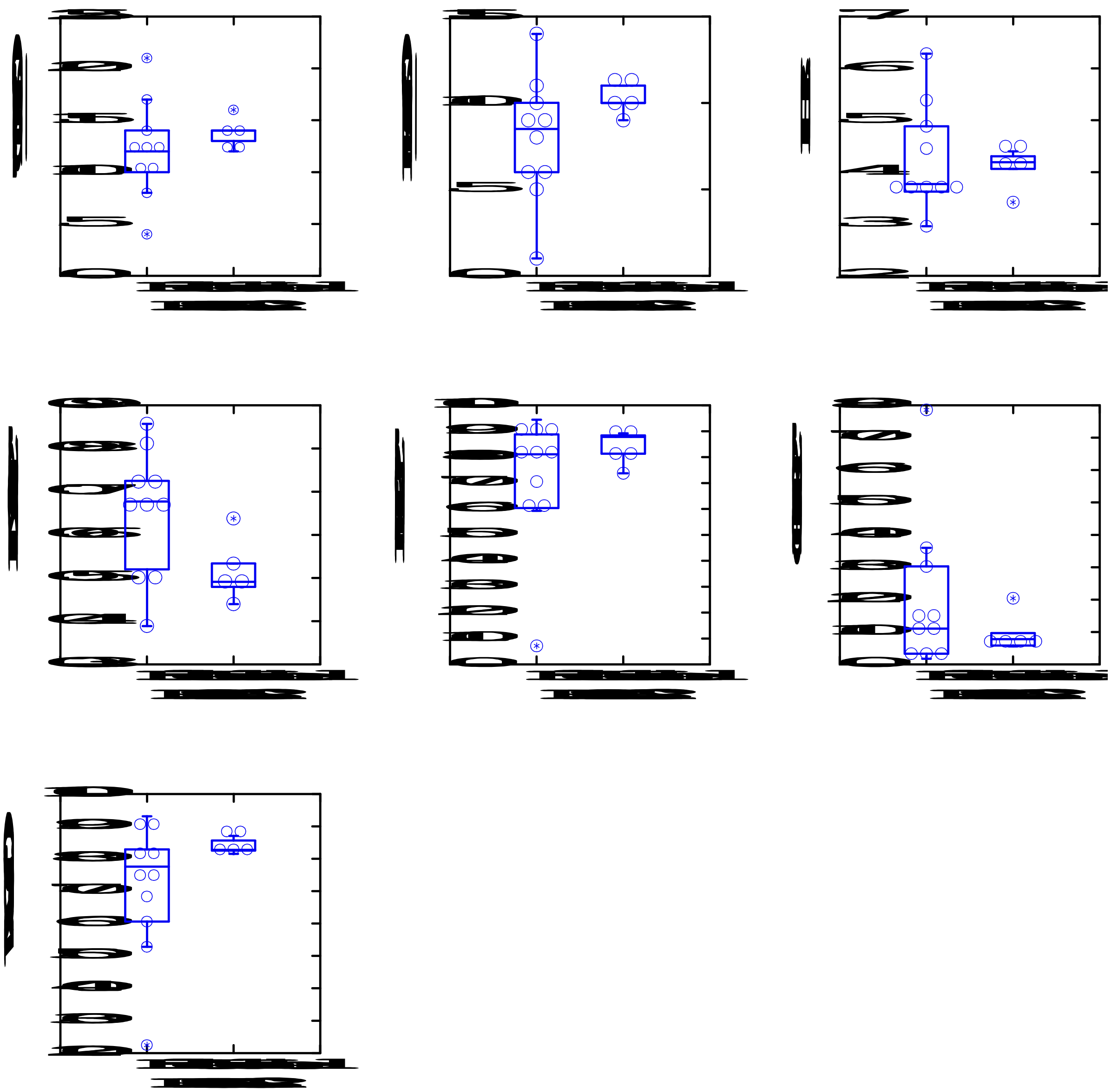


Figure C-5. Box plots of the WVSCI and its component metrics versus the EIS class for the autumn 2000 season. Circles represent site scores.

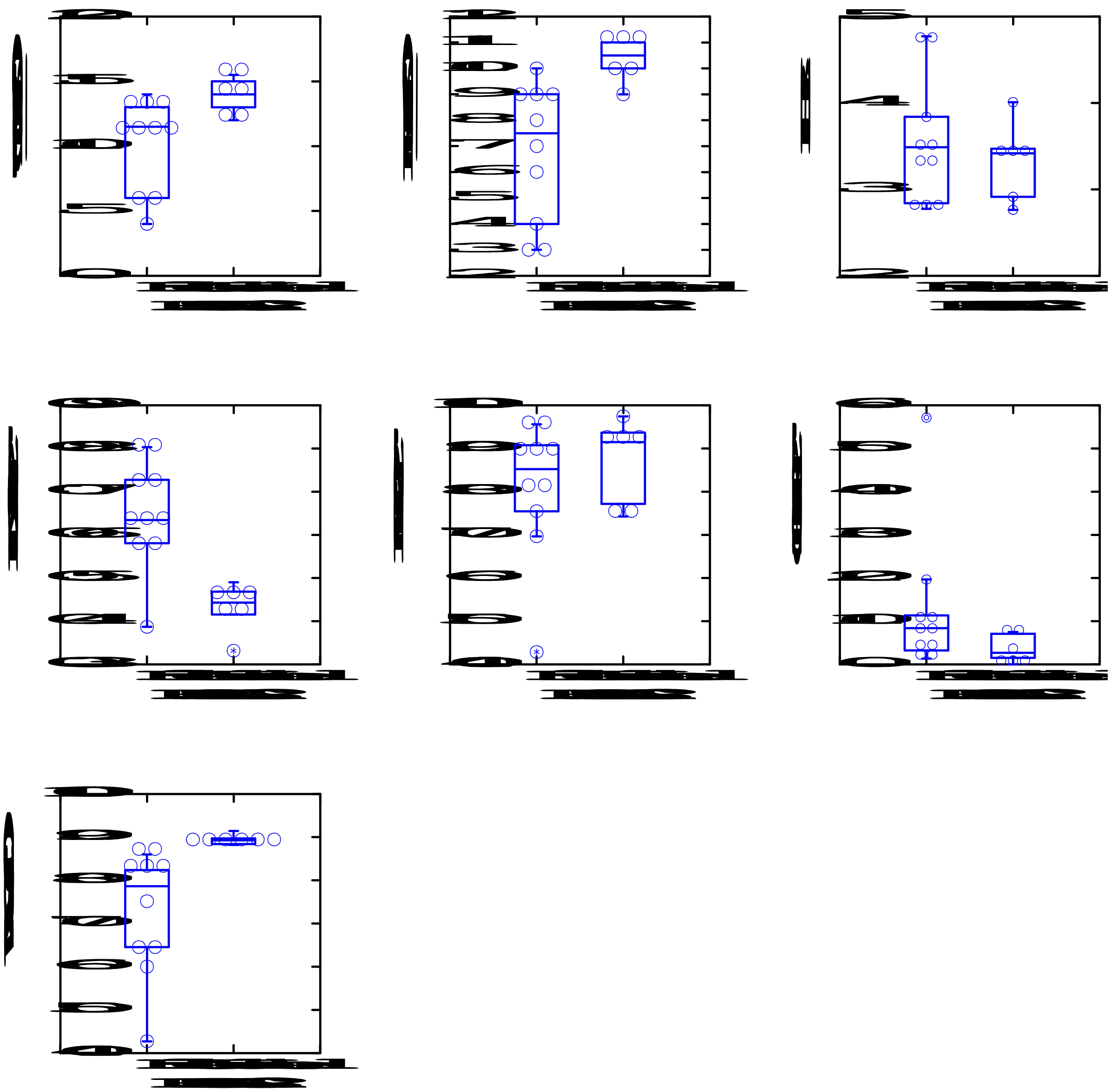


Figure C-6. Box plots of the WVSCI and its component metrics versus the EIS class for the winter 2001 season. Circles represent site scores.

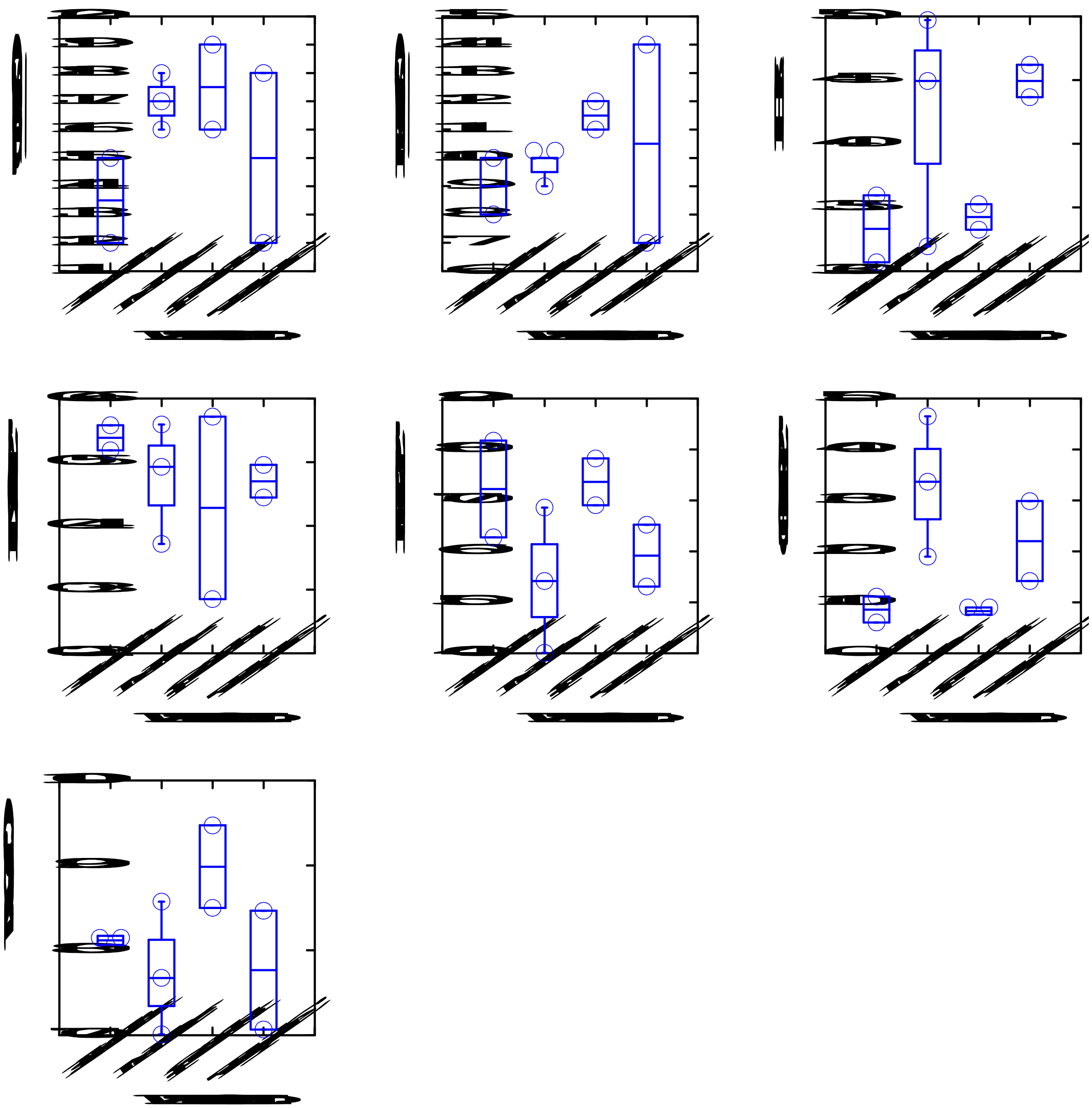


Figure C-7. Box plots of the WVSCI and its component metrics versus watershed for unmined sites in the spring 1999 season.

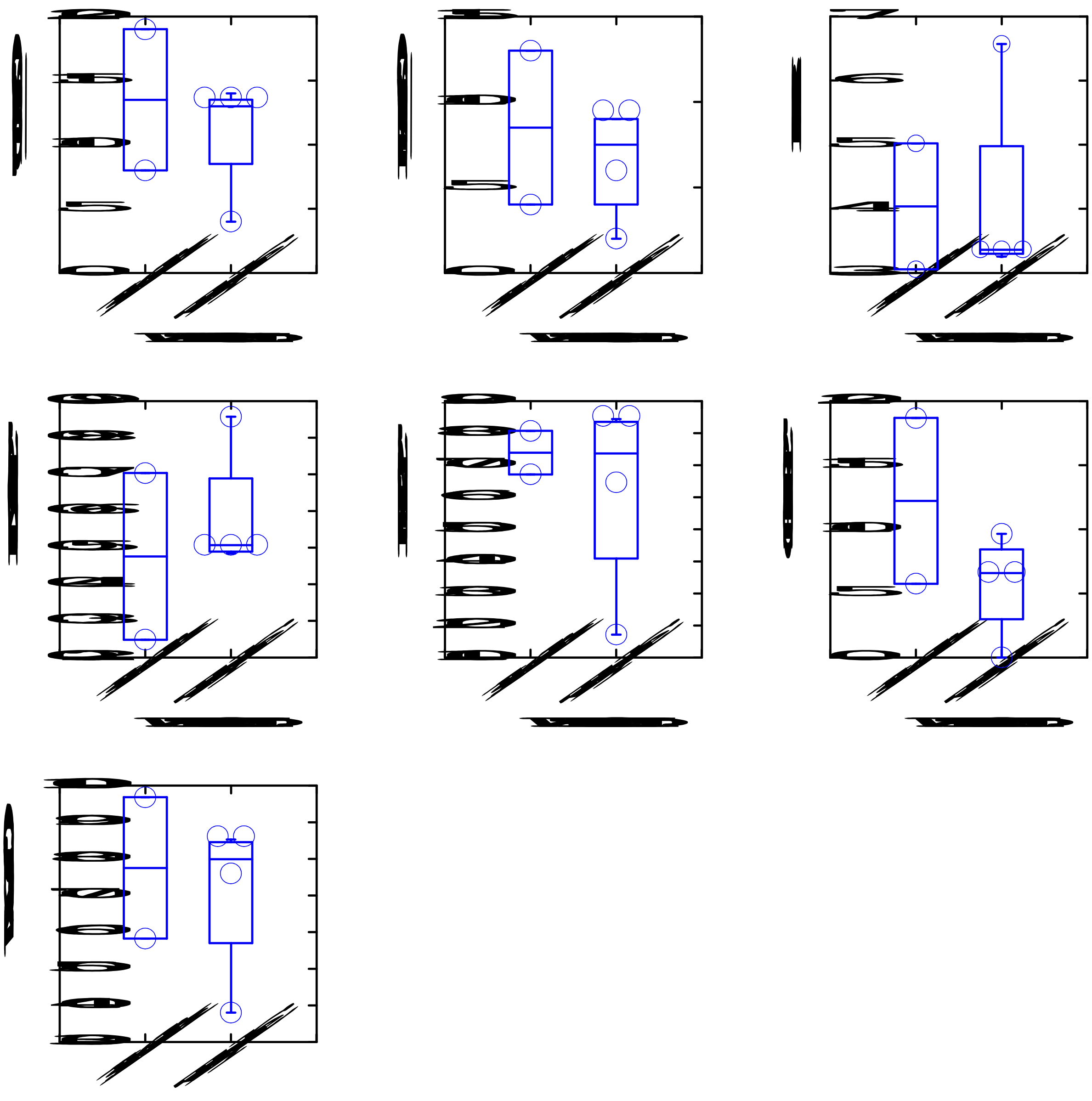


Figure C-8. Box plots of the WVSCI and its component metrics versus watershed for unmined sites in the autumn 1999 season.

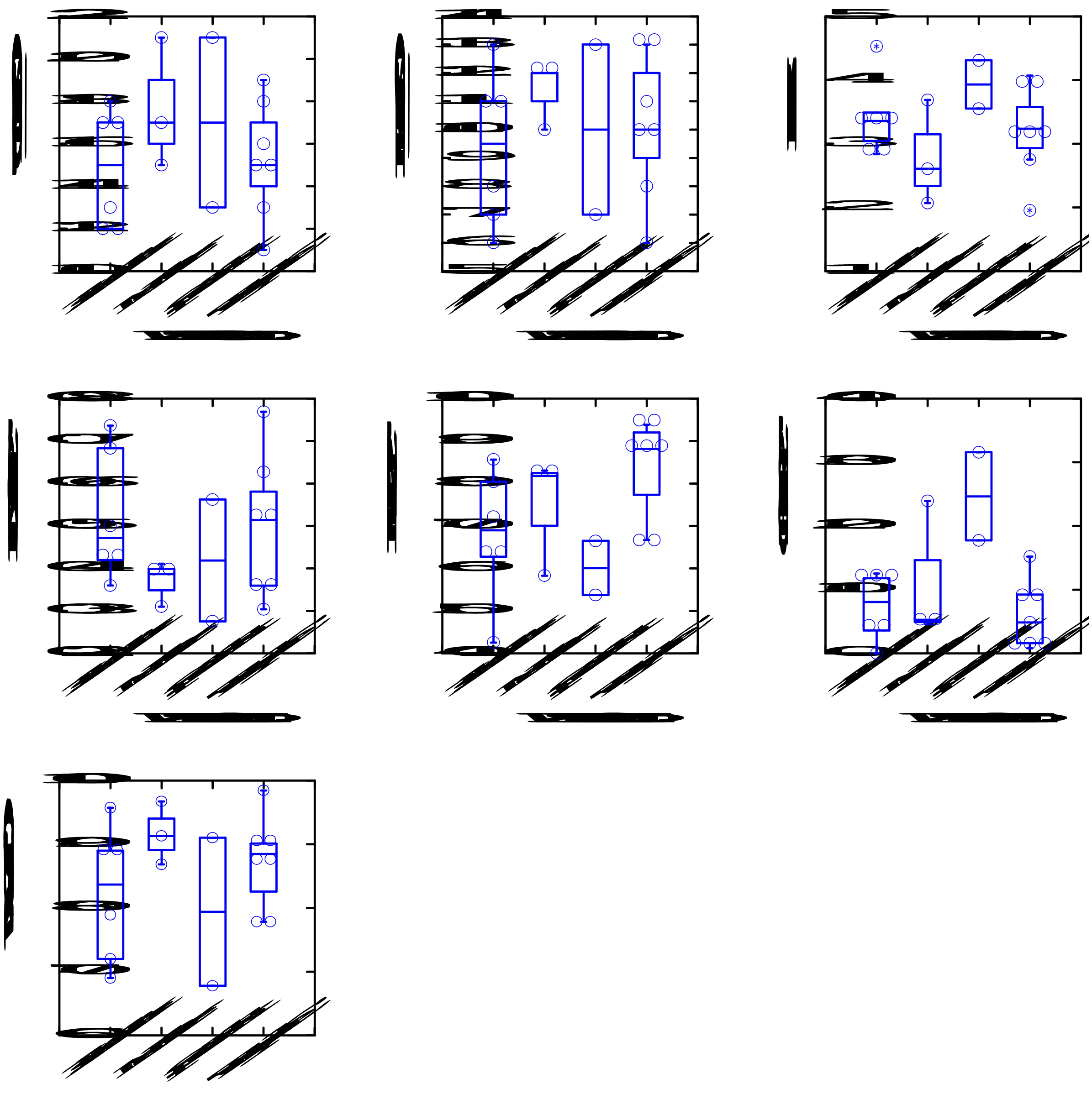


Figure C-9. Box plots of the WVSCI and its component metrics versus watershed for unmined sites in the winter 2000 season.

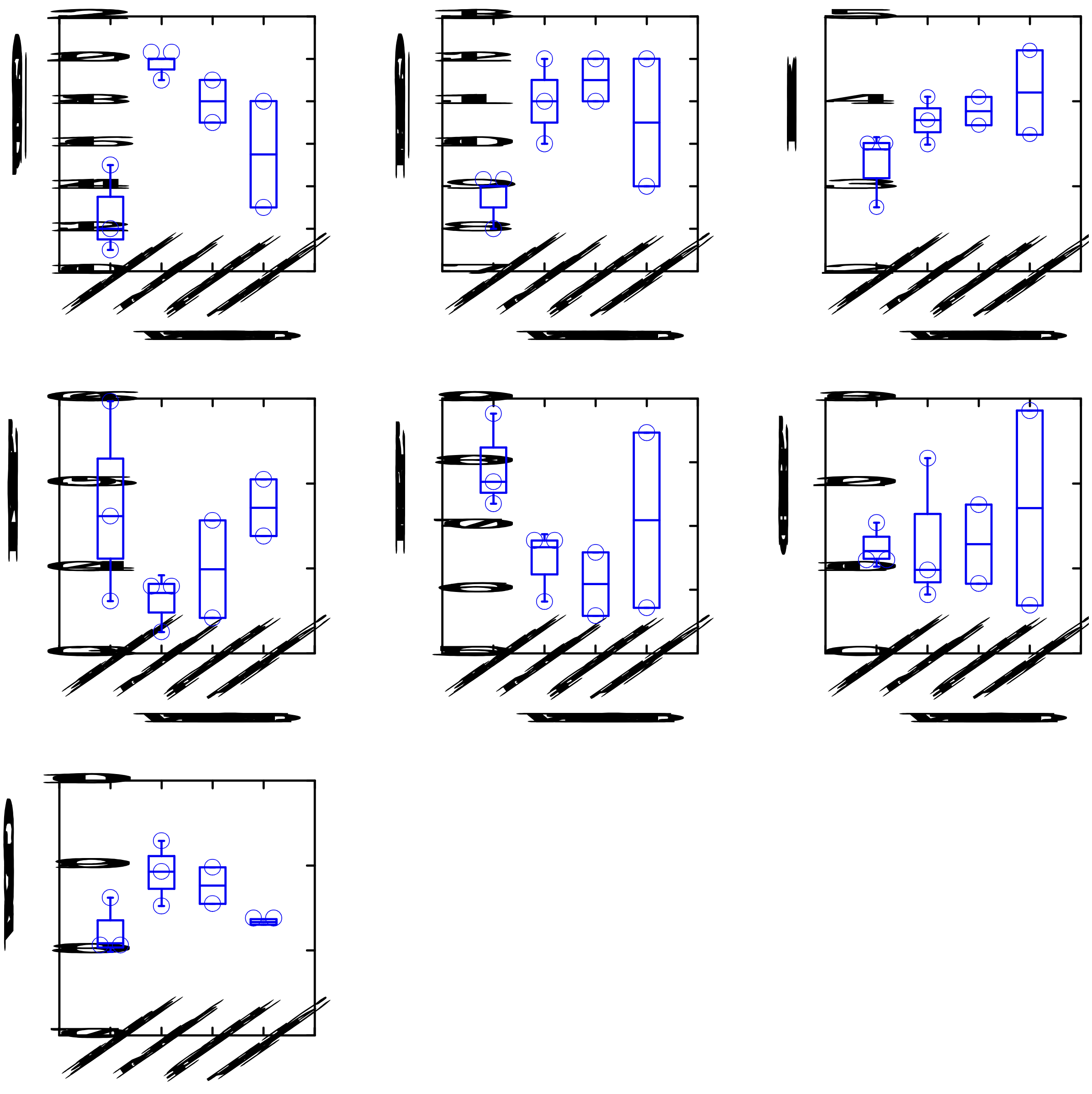


Figure C-10. Box plots of the WVSCI and its component metrics versus watershed for unmined sites in the spring 2000 season.

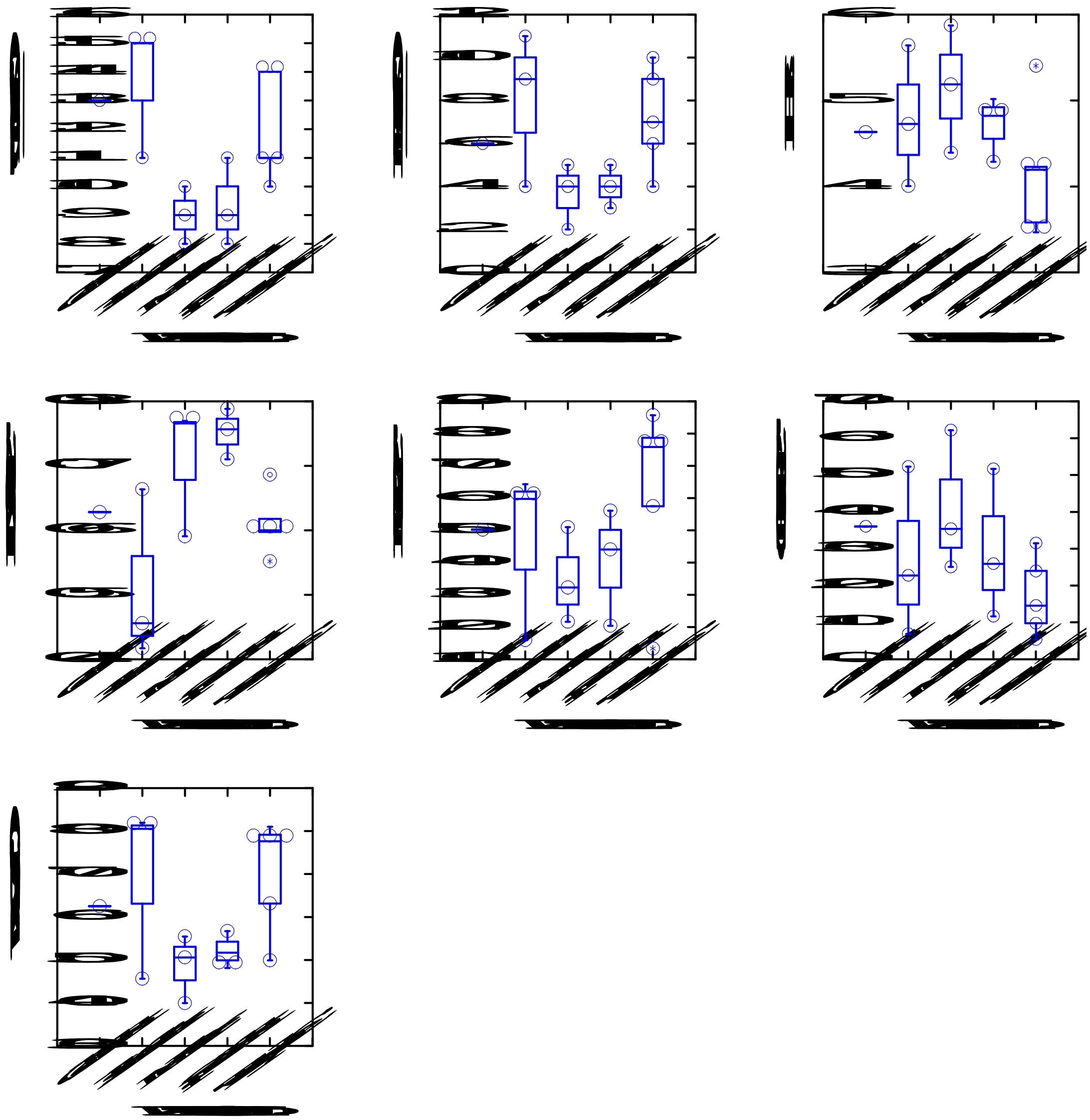


Figure C-11. Box plots of the WVSCI and its component metrics versus watershed for Filled sites in the spring 1999 season. Circles represent site scores.

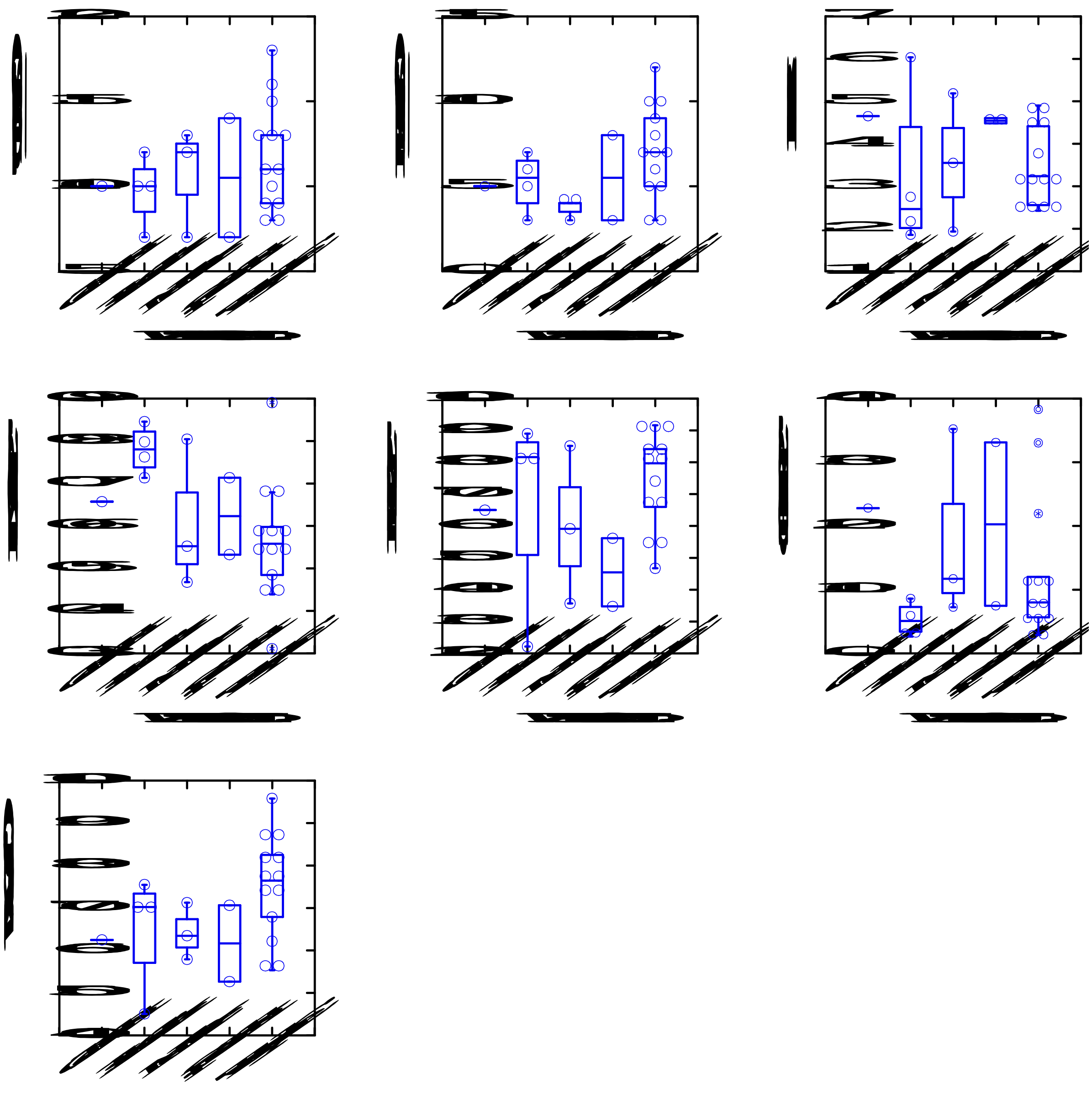


Figure C-12. Box plots of the WVSCI and its component metrics versus watershed for Filled sites in the autumn 1999 season. Circles represent site scores.

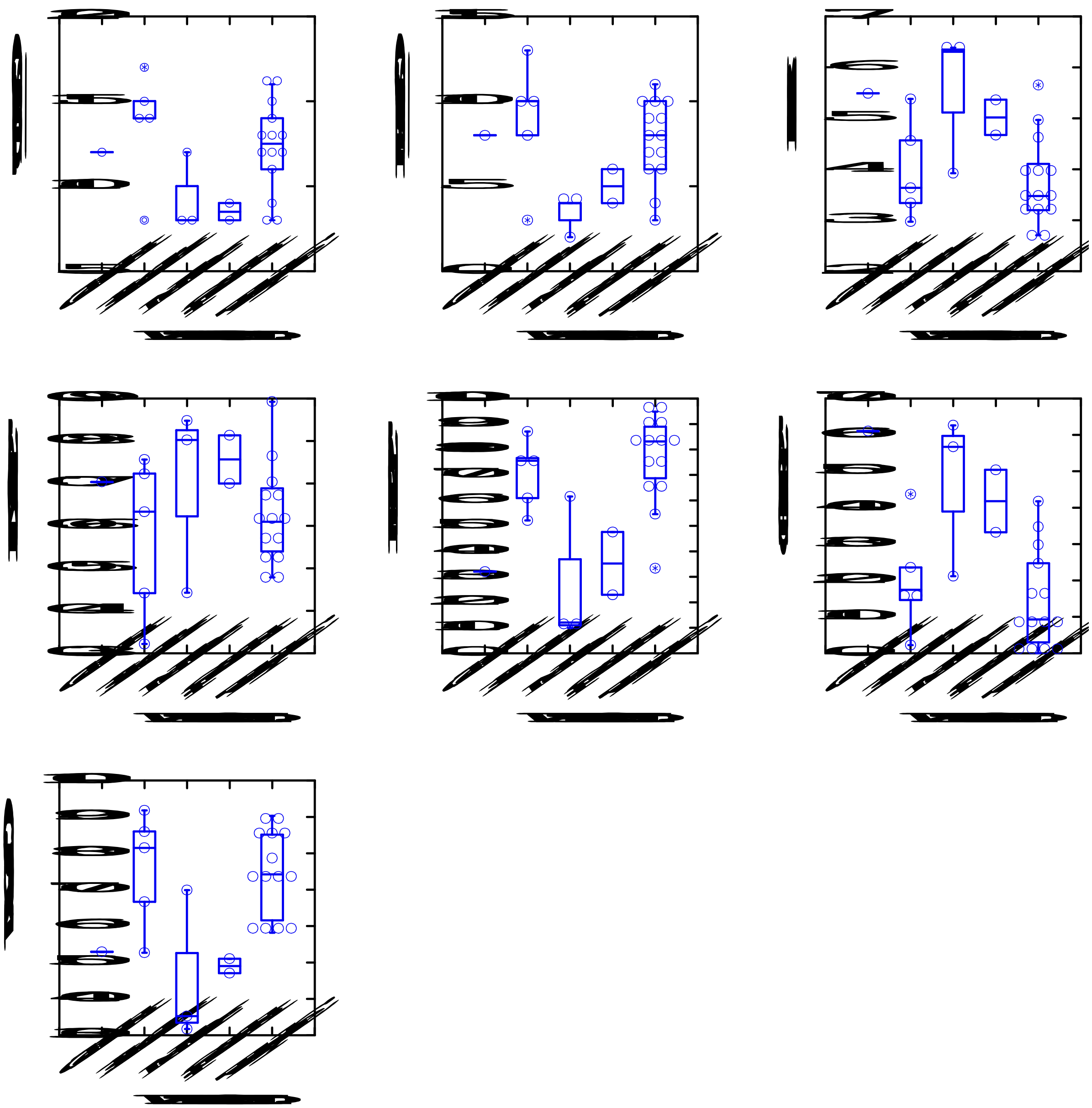


Figure C-13. Box plots of the WVSCI and its component metrics versus watershed for Filled sites in the winter 2000 season. Circles represent site scores.

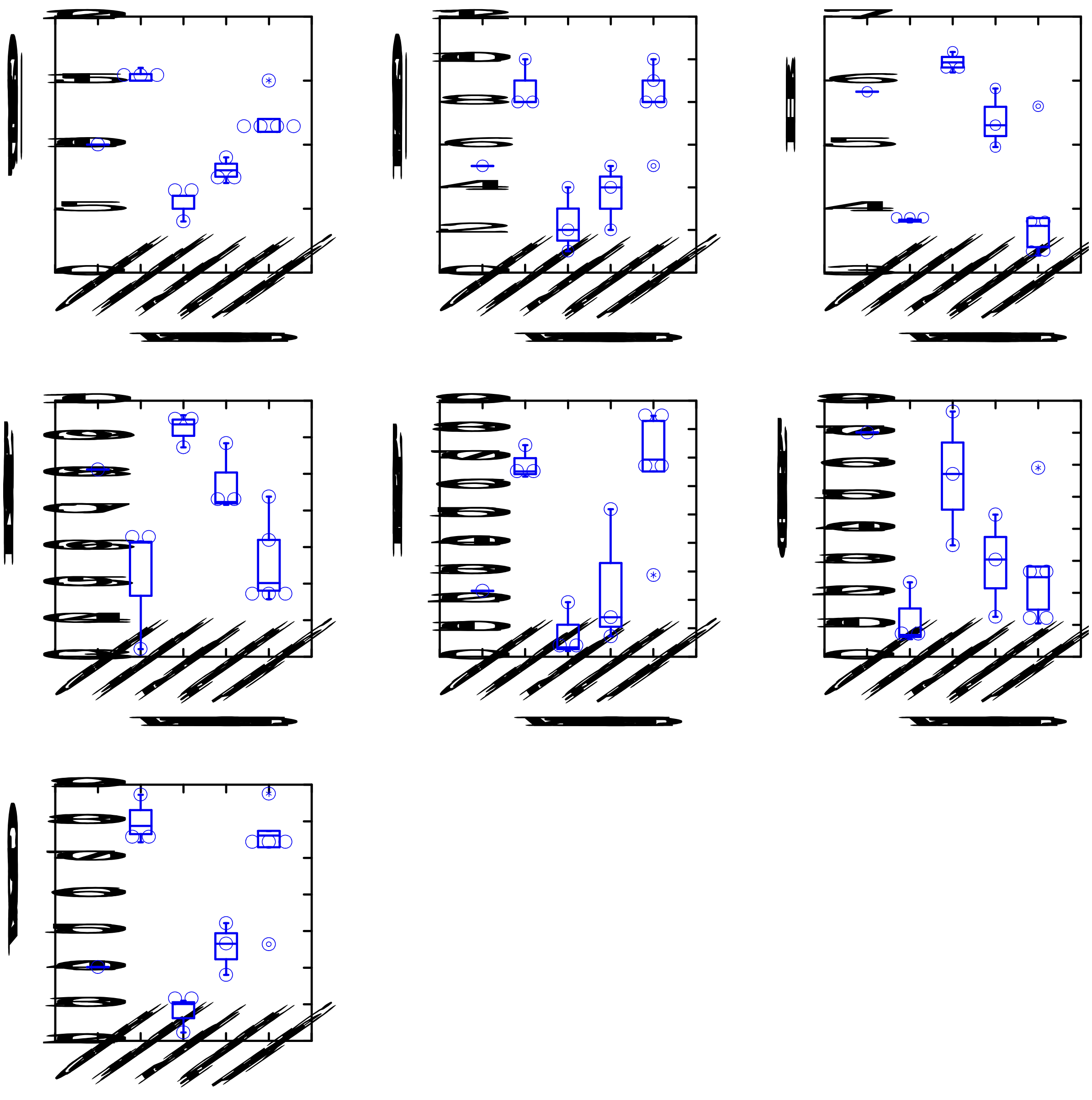


Figure C-14. Box plots of the WVSCI and its component metrics versus watershed for Filled sites in the spring 2000 season. Circles represent site scores.

APPENDIX D

SCATTER PLOTS OF THE WVSCI VERSUS KEY WATER QUALITY PARAMETERS

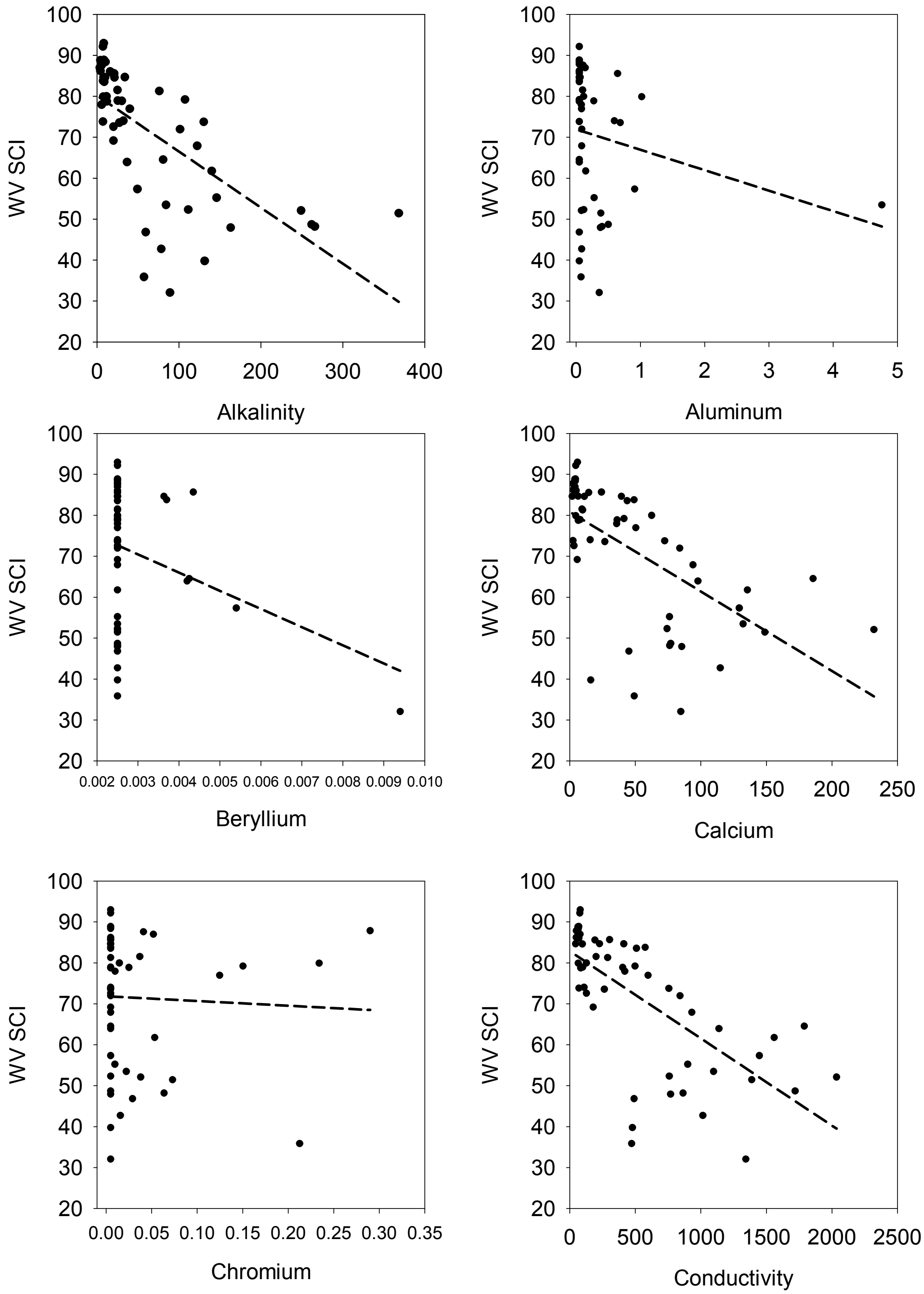

WV SCI
Alkalinity
Aluminum
Beryllium
Calcium
Chromium
Conductivity

Figure D-1. The WVSCI, rarefied to 100 organisms, versus water quality parameters. Dashed line represents best fit line using linear regression.

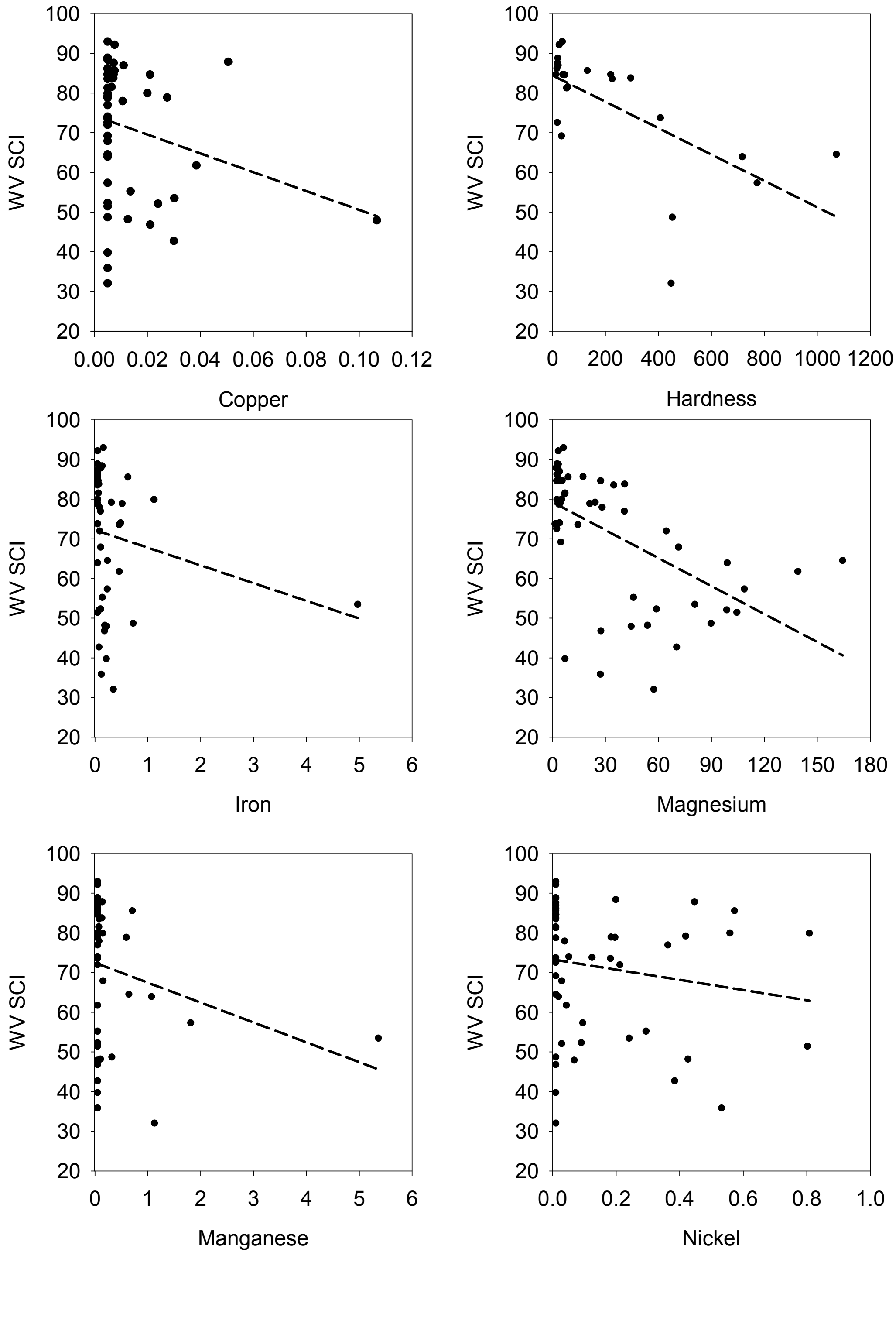

WV SCI
Copper
Hardness
Iron
Magnesium
Manganese
Nickel

Figure D-1. Continued.

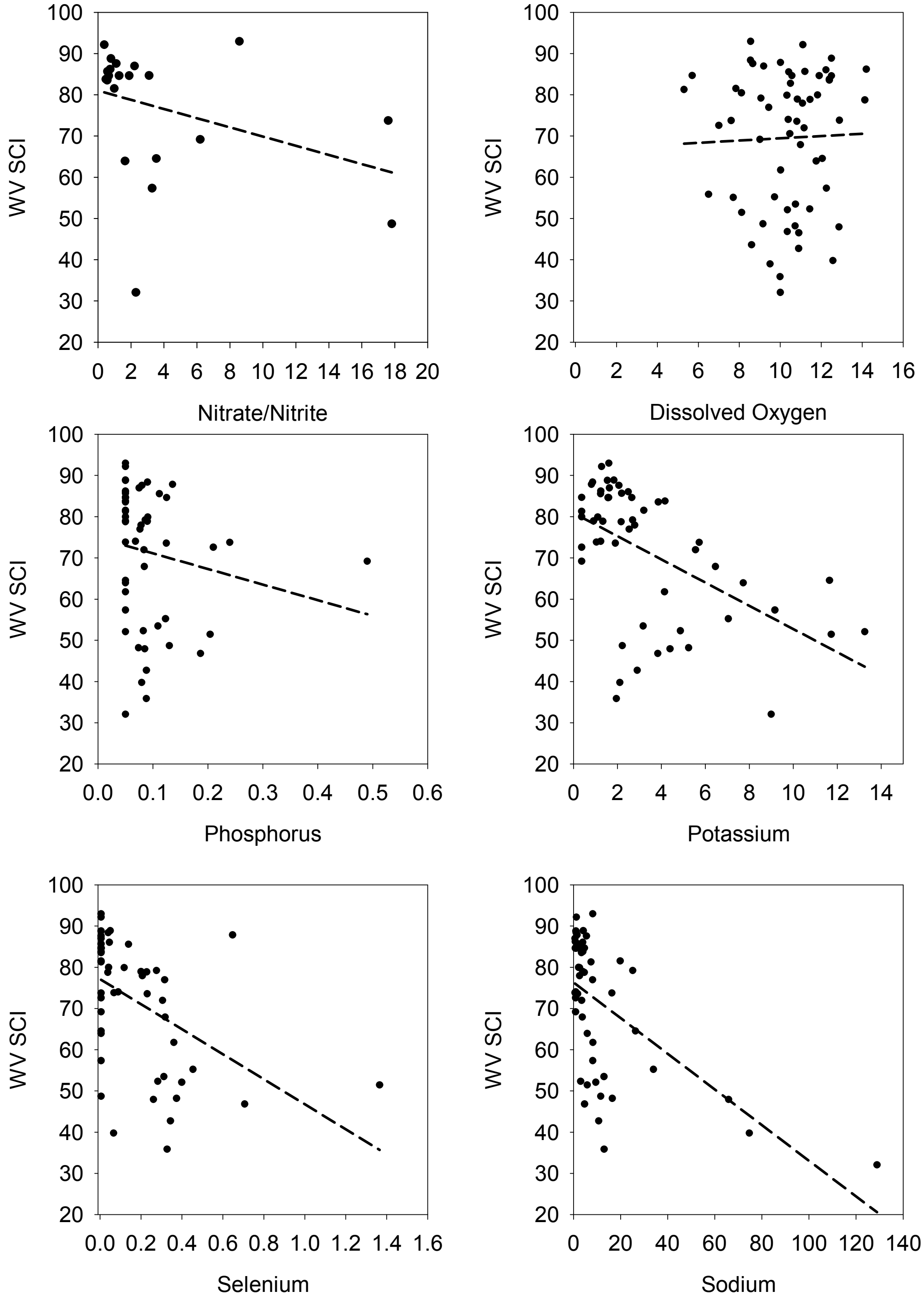


Figure D-1. Continued.

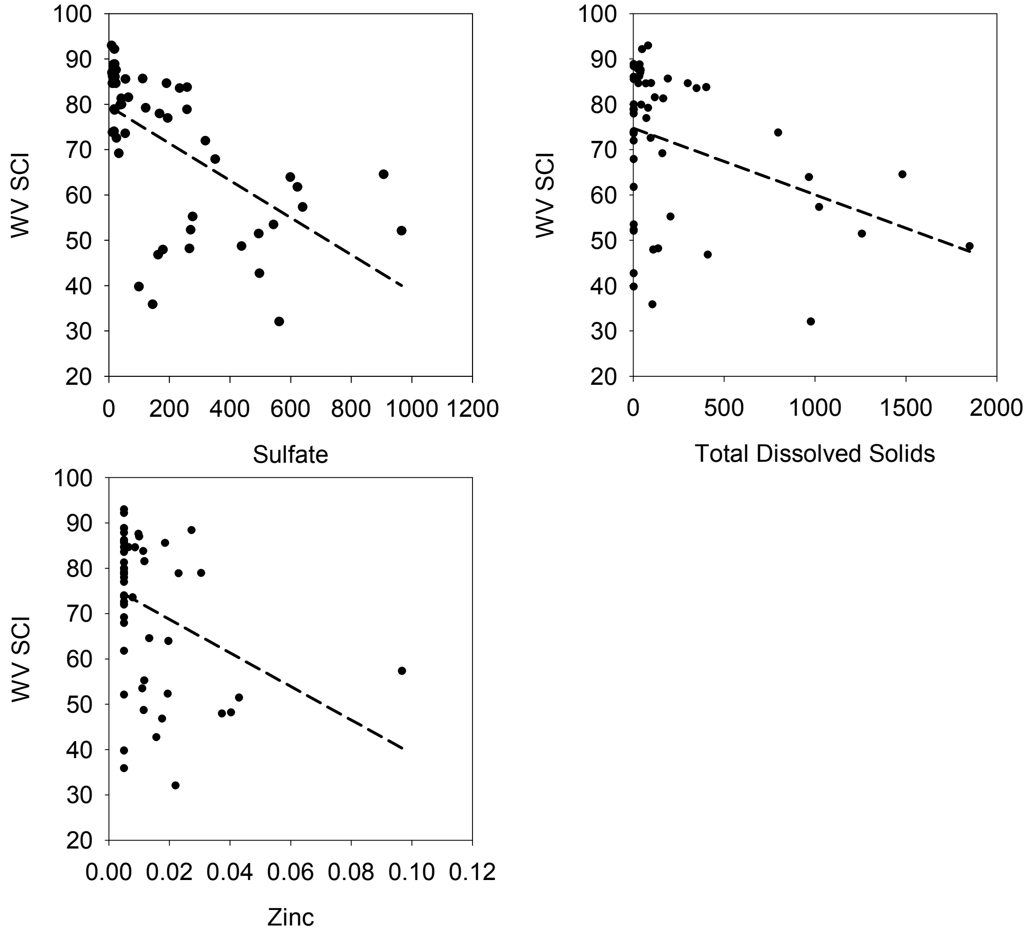


Figure D-1. Continued.

APPENDIX E
STANDARDIZATION OF DATA AND METRIC CALCULATIONS

Standardization and Statistical Treatment of MTM/VF Fish Data

Fish Sample Collection Methods

Fish communities, like benthic communities, respond to changes in their environment. Some fish species are less tolerant of degraded conditions; as stream health decreases, they will either swim away or perish. Other species are more tolerant of degraded conditions, and will dominate the fish community as stream health declines.

Fish are collected using a backpack electrofisher. In electrofishing a sample area, or “reach”, is selected so that a natural barrier (or a block net, in the absence of a natural barrier) prevents fish from swimming away upstream or downstream. An electrical current is then discharged into the water. Stunned fish float to the surface and are captured by a net, and held in buckets filled with stream water. The fish are identified, counted and often measured and/or weighed. Three passes are made with the electrofisher to collect all the fish in the selected stream reach. After the three passes are complete and the fishes have recovered, they are released back to their original habitat. Some fish may be retained as voucher specimens. The data collected from the three passes are composited into a single sample for the purposes of the MTM-VF project.

Pennsylvania State University (PSU) conducted fish sampling for USEPA. PSU collected fish from 58 sites located on first through fifth order streams in West Virginia. Fish were also sampled by REIC, Potesta, and BMI, following the same protocols. The only exceptions were five samples taken by REIC that were made with a pram electrofisher. In a pram unit, the electrofishing unit is floated on a tote barge rather than carried in a backpack. Otherwise, the pram samples followed the same protocols.

The Mid-Atlantic Highland IBI

The Mid-Atlantic Highland Index of Biotic Integrity, or IBI, (McCormick et al. 2001), provides a framework for assessing the health of the fish community, which, like the WV SCI, indicates the overall health of a stream. The IBI was developed and calibrated for the Mid-Atlantic Highlands using samples from several Mid-Atlantic states, including West Virginia. The IBI is a compilation of scores from nine metrics that are responsive to stress (Table E-1).

Table E-1. Metrics included in the Mid-Atlantic Highland IBI, with descriptions and expected response to increasing degrees of stress.

Metric	*Metric Description*	*Predicted Response to Stress*
Native Intolerant Taxa	**Number of indigenous taxa that are sensitive to pollution; adjusted for drainage area**	**Decrease**
Native Cyprinidae Taxa	**Number of indigenous taxa in the family Cyprinidae (carps and minnows); adjusted for drainage area**	**Decrease**
Native Benthic Invertivores	**Number of indigenous bottom dwelling taxa that consume invertebrates; adjusted for drainage area**	**Decrease**
Percent Cottidae	**Percent individuals of the family Cottidae (sculpins)**	**Decrease**
Percent Gravel Spawners	**Percent individuals that require clean gravel for reproductive success**	**Decrease**
Percent Piscivore/Invertivores	**Percent individuals that consume fish or invertebrates**	**Decrease**
Percent Macro Omnivore	**Percent individuals that are large and omnivorous**	**Increase**
Percent Tolerant	**Percent individuals that are tolerant of pollution**	**Increase**
Percent Exotic	**Percent individuals that are not indigenous**	**Increase**

Watershed Standardization

In nature, larger watersheds are naturally more diverse than smaller watersheds. Not surprisingly, this was found to be true in the MTM-VF project. To ensure that differences among fish communities are due to differences in stream health and not from the natural effect of watershed size, three richness metrics were standardized to a 100km^2 watershed.
This standardization applies only to the three richness metrics; percentage metrics are not affected by watershed size and required no adjustment before scoring.

The regression equations used in the watershed standardization were developed by McCormick et al. 2001. They studied the relationship between watershed size and fish community richness in minimally stressed sites, and derived equations that predict the number of taxa that would be expected in a healthy stream of a given watershed size. The equations were not published in the original 2001 paper, but were obtained from McCormick in a personal communication.

First, the predicted numbers of taxa were calculated using the regression equations. Then residual differences were calculated:

Residual difference = Actual number in sample – Predicted number

Finally, an adjustment factor was added to the residual difference (see Table E-2), depending on the richness metric.

Table E-2. Regression equations and adjustment factors for standardizing richness metrics to a 100 km2 watershed. (McCormick, personal communication)

Richness Metric	*Regression Equation*	*Adjustment Factor*
Native Intolerant Taxa	**predicted = 0.440071 + 0.515214 * Log_{10} (Drainage Area [km^2])**	**1.470**
Native Cyprinidae Taxa	**predicted = 0.306788 + 2.990011 * Log_{10} (Drainage Area [km^2])**	**6.287**
Native Benthic Invertivores	**predicted = 0.037392 + 2.620796 * Log_{10} (Drainage Area [km^2])**	**5.279**

Metric Scoring and IBI Calculation

After the necessary watershed adjustments had been made, metric scores were applied to the adjusted richness metrics and the raw percentage metrics. The scoring regime was originally derived from the distribution characteristics of the large Mid-Atlantic Highlands data set upon which the IBI was calibrated (McCormick et al. 2001).

Some metrics decrease in value with increasing stress, such as the richness metrics. For example, the number of intolerant species (those sensitive to poor water quality) decreases as stream health declines. Each of the metrics that decreases in value with increasing stress was given a score ranging from 0 – 10 points. Zero points were given if the adjusted value was less than the 5th percentile of McCormick's non-reference sites; 10 points were given if the adjusted value was greater than the 50th percentile of McCormick's high quality reference sites. Intermediate metric values, those between 0 and 10, were interpolated between the two end points.

Other metrics increase in value with increasing stress, such as the percent of tolerant fish species. As stream health declines, only the tolerant species thrive. Metrics that increase in value with increasing stress are also given a score ranging from 0 to 10. A score of 0 points is given to values greater than the 90th percentile of McCormick's non-reference sites. A score of 10 points are given to values less than the 50th percentile of McCormick's moderately restrictive reference sites. Intermediate metric values were scored by interpolation between 0 and 10.

After all nine metrics have been scored, they are summed. Nine metrics scoring a possible 10 points each equals a possible maximum of 90 points; to convert to a more easily understood 100-point scale, the raw sum score is multiplied by 1.11. The Mid-Atlantic Highlands IBI is this resulting number, on a scale of 0-100 (Table E-3).

Table E-3. Mid-Atlantic Highland IBI: Metric scoring formulas. Richness metrics were adjusted for drainage area before calculating scores.

Metric	***Scoring formulas (X=metric value)***
Native Intolerant Taxa (Adjusted for watershed)	**If X>1.51, then 10. If X<0.12, then 0. Else 10*X/1.39**
Native Cyprinidae Taxa (Adjusted for watershed)	**If X>6.24, then 10. If X<1.54, then 0. Else 10*X/4.70**
Native Benthic Invertivore Taxa (adjusted for watershed)	**If X>5.34, then 10. If X<1.27, then 0. Else 10*X/4.07**
Percent Cottidae	**If X>7, then 10. Else 10*X/7**
Percent Gravel Spawners	**If X>72, then 10. If X<21.5, then 0. Else 10*X/50.5**
Percent Piscivore/Invertivores	**If X>9, then 10. Else 10*X/9**
Percent Macro Omnivore	**If X>16, then 0. If X<0.2, then 10. Else 10*(16-X)/15.8**
Percent Tolerant	**If X>97, then 0. If X<28, then 10. Else 10*(97-X)/69**
Percent Exotic	**If X>24, then 0. If X<0.2, then 10. Else 10*(24-X)/23.8**
SUM of all 9 metric scores	**Raw Score**
Mid-Atlantic Highland IBI score (0-100 range)	**Raw Score x 1.11**

Standardization and Metric Calculations of Benthic Data

Benthic Sample Collection Methods

What do we know about healthy Appalachian streams? There are many species of organisms that live in streams (insects, crustaceans, mussels, worms), and in general, healthy streams have a greater variety of animals than unhealthy streams. Three groups of insects in particular, the mayflies, stoneflies, and caddisflies, are sensitive to pollution and degradation and tend to disappear as a stream's water quality decreases. Other insect groups are more tolerant to pollution, and tend to increase as a percentage of the total benthic (bottom-dwelling) communities in unhealthy streams. In order to determine whether a stream is healthy or unhealthy, we must obtain a representative estimate of the variety and identity of species in the stream.

How do biologists sample stream communities to get a representative and precise estimate of the number of species? First, we must know where the organisms live in the stream. An Appalachian stream bottom is not a uniform habitat: there are large rocks, cobble, gravel, patches of sand, and tree trunks in the streambed. Each of these is a microhabitat and attracts species specialized to live in the microhabitat. For example, some species live on the tops of rocks, in the current, to catch food particles as they drift by. Some species crawl around in protected areas on the underside of rocks; some cling to fallen tree trunks or branches; yet others live in gravel or sand. Clearly, if we sample many microhabitats, we will find more species than if we sample only one. In order to characterize the stream section, we need to sample a large enough area to ensure that we have sampled most of the microhabitats present.

How do we “measure” the biological effects of human activities, such as mining, on stream ecosystems? What is the unit of the stream that we characterize? Typically, we wish to know the effects on a wide variety of organisms throughout the stream. However, sampling everything is expensive and potentially destructive. Selecting a single, common habitat that is an indicator of stream condition is analogous to a physician measuring fever with an oral thermometer at a single place (the mouth). Therefore, biologists selectively sample riffles, which are prevalent in Appalachian streams, and are preferred habitat for many sensitive species. When we sample a riffle, we wish to characterize the entire riffle, not just an individual rock or patch of sand, and sampling must represent the microhabitats present. By taking several samples, even with a relatively small sampling device such as a Surber Sampler, we can ensure that enough microhabitats have been sampled to obtain an accurate estimate of diversity in the stream.

Sampling Gear

Sampling also depends on the gear and equipment that biologists use to capture organisms. Small samplers and nets can be easily and economically handled by one or two persons; larger sampling equipment requires larger crews. In the MTM-VF project, the sampling protocol calls for 6 Surber samples (0.09 square meter each, for 0.56 square meter total from each site), or 4 D-frame samples (0.25 square meter each, for 1 square meter from each site). If the Surber or D-frame grabs are spread out throughout the riffle (preferably in a random manner), then they will adequately represent most of the microhabitats present, and total diversity of the riffle can be characterized.

Standardization of data

Many agencies were involved in the collection of data for the Mountain Top Mining Environmental Impact Statement. Not all organizations used the same field sampling methods, and during the two-year investigation, some organizations changed their sampling methods. In order to "compare apples to apples," it is necessary to standardize the data, so that duplicate samples taken using different methods will yield the same results after standardization.

We begin here with a description of the sampling methods used, a general discussion of sampling, analysis of a set of paired samples using two methods, and finally the specific steps used to standardize the samples from the different organizations.

MTM/VF Benthic Sampling Methods

The two methods used in the MTM/VF study, which we term the "D-frame method" and the "Surber method," differ in sampling gear and in the treatment of the collected material. The methods are compared below.

D-frame Method

Equipment: A D-frame net is a framed net, in the shape of a "D", which is attached to a pole.

Procedure: The field biologist positions the D-frame net on the stream bottom, then dislodges the stream bottom directly upstream to collect the stream-bottom material, including sticks and leaves, and all the benthic organisms. The net is 0.5 meter wide, and $0.25m^2$ area of streambed is sampled with each deployment. In the MTM/VF study, the net was deployed 4 times at each site, for a total area of 1.0 m^2.

Compositing: All the collected materials were composited into a single sample.

Subsampling: Samples collected in the D-frame method are often quite large, and two organizations "subsampled" to reduce laboratory processing costs. In subsampling, the samples are split using a sample splitter (grid), and a subsample consisting of 1/8th (or, in the case of samples with few organisms, 1/4th or 1/2) of the original material was analyzed. All organisms in the subsample were identified and counted.

Surber Method

Equipment: A Surber sampler is a square frame, covering 1 square foot ($0.093m^2$) of stream bottom.

Procedure: The Surber is placed horizontally on cobble substrate in shallow stream riffles. A vertical section of the frame has the net attached and captures the dislodged organisms from the sampling area.
In the MTM/VF study, the Surber sampler was deployed 3 to 6 times at each site, for a total area sampled of 3 to 6 square feet (0.28 to $0.56m^2$).

Compositing: The materials collected were not composited, but were maintained as discrete sample replicates.

Subsampling: The materials collected in each of the Surbers were not subsampled. All organisms were identified and counted.

The D-frame sampler was most consistently used by participants. EPA and Potesta used only D-frame sampling; BMI used only D-frame sampling in the first two sets of samples, and afterwards used both Surber and D-frame samplers. REIC collected both Surber and D-frame samples throughout the study. The various methods used by the organizations participating in the MTM/VF study are summarized in Table E-4.

Table E-4. A comparison of each organization's methods of collecting and compositing samples, and laboratory subsampling protocols.

Organization	*Sample Method*	*Compositing*	*Subsampling*
USEPA	4 times 1/4m^2 D-frame net	Composited samples	1/8 of original sample. If abundance was low, the laboratory subsampled to 1/4 or ½ of the original sample, or did not subsample at all.
REIC (Twelvepole Creek)	3 times Surber and 4 times 1/4m^2 D-frame net	All Surber samples were analyzed separately (no compositing). Composited samples.	The D-frame samples were subsampled to 1/4 of original sample if necessary. All 7 samples were combined for reporting, representing approximately 1.3 m^2 of stream bottom.
Potesta (Twenty Mile Creek)	4 times 1/4 m^2 D-frame net.	Composited samples	Not subsampled; counted to completion.
BMI (Twenty Mile Creek)	Fall 1999 and Spring 2000: 4 times 1/4 m^2 D-frame net. Fall 2000, 6 times Surber, and four times 1/4 m^2 D-frame net. Spring 2001, 4 times Surber and four times 1/4m^2 D-frame sample**.**	Composited samples. Surber samples kept separate. D-frame samples were composited. Surber samples kept separate. D-frame samples were composited.	Not subsampled; counted to completion. Not subsampled; counted to completion. Not subsampled; counted to completion.
BMI (Island Creek):	Fall 1999 and Spring 2000, four times 1/4 m^2 D-frame net, Fall 2000, 4 times Surber, kept separate, and four times 1/4 m^2 D-frame net, composited. Spring 2001: No data.	Composited samples. Surber samples were kept separate. D-frame samples were composited.	Not subsampled; counted to completion. Not subsampled; counted to completion.

Treatment of Sampler Data

How do we treat data from the samplers? A common method is to take the average of measures from several (4 or 6) samplers. The problem with this approach is that we know that each sampler, individually, underestimates species richness of the stream site; thus the average of underestimates will also be an underestimate (see Table E-5). In addition to species (or family) richness, a measure important in the West Virginia Stream Condition Index, and in many other similar condition indexes, is the degree to which a community is dominated by the most abundant species found. In degraded streams, communities are often dominated by one or a few species tolerant of poor habitat or poor water quality. In a healthy stream, dominance over the entire community is low. However, a single microhabitat, such as a large rock, is likely to by dominated by one or two species adapted to that microhabitat. A different species will be dominant in a sand habitat. The entire riffle is diverse and has low dominance when we consider several microhabitats. Thus, if we calculate the average dominance over several small sampling devices, such as Surbers, we overestimate community dominance. Each Surber sample may be highly dominated by a different species, yet the overall community may not dominated by any of those species. This is shown with data from one of the sites (Table E-5): average richness of Surbers is lower than richness of the composited Surbers (representing the entire riffle). Average dominance of the Surbers is higher than the composited sample. By averaging, this site appears to be in poorer condition than it really is, especially if compared to West Virginia's Stream Condition Index.

Standardizing Sampling Effort

Sampling effort is a combination of the total riffle area sampled, the heterogeneity of the stream bottom sampled, and the number of organisms identified. As previously discussed, a composited sample that consists of several smaller samples from throughout the riffle area will adequately characterize the abundances and relative abundances of most of the common species at a site. It will not, however, necessarily characterize all of the rare species at a site (those making up less than about 2% of the total community). Sampling to collect all rare species is prohibitively expensive and destructive of the riffle. But we must consider the effects of rare species since they contribute to diversity and richness measures in proportion to sampling effort. For example, the D-frame net, which covers 1 m^2, (10.8 square feet) will capture more rare species than 4 or 6 Surber samplers, which cover only 0.37 m^2 (4 square feet) and 0.56 m^2 (6 square feet) respectively. By the same token, subsampling, or counting only a portion of the total sample, also undercounts rare species.

Fortunately, it is relatively easy to standardize sampling effort among different sampling methods so that the bias is removed. Standardization is done by adjusting taxa counts to expected values for subsamples smaller than an original sample, using the following binomial probabilities for the capture of each taxon (Hurlbert 1971; Vinson and Hawkins 1996).

$$E(S_n) = \sum_i \left[1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right]$$

= The expected number of species in a sample of n individuals selected at random from a collection containing *N* individuals, *S* species, and N_i individuals in the *i*th species.

Taxa counts (number of species or families) can only be adjusted down to the level of the smallest sampling effort in the data set; it is not possible to estimate upwards (and effectively "make up" data). In the MTM/VF data, benthic samples were standardized to 200 individuals, which is the standard WV SCI practice, and to 100 individuals, to accommodate those samples that contained less than 200 organisms. Individual taxa are not removed from a sample in the standardization process; only the taxa counts are standardized. Estimates of abundance per area and relative abundance are unaffected by sampling effort, and are not adjusted.

Table E-5. Six Surber replicates from site MT-52 (Island Creek), Fall 1999. The dominant family for each Surber is in bold, outlined with a heavy line. The subdominant family is outlined with a light line. Either Taeniopterygidae or Nemouridae are dominant in each Surber, but they tend not to co-occur in the same Surber. Metrics are shown at the bottom.

			Surber					
Order and family	A	B	C	D	E	F	Composite	
Beetles								
Elmidae	11	13	3	3	14		44	
Psephenidae	6	2	4	4	9		25	
Caddisflies								
Hydropsychidae	13		4	6	8	11	42	
Philopotamidae			1	2			3	
Polycentropodidae				8	5		13	
Rhyacophiloidea	8	8	4			6	26	
Uenoidae	1	2			5	3	11	
Mayflies								
Ameletidae	11		1			19	31	
Baetidae		3	1	5	18		27	
Baetiscidae	1						1	
Ephemerellidae	3	6	4	3	16	10	42	
Heptageniidae		2					2	
Stoneflies								
Chloroperlidae	1					1	2	
Nemouridae	**50**		**61**			**24**	135	
Perlidae		1					1	
Perlodidae		23		1			24	
Taeniopterygidae		**71**	1	**25**	**95**		**192**	
True flies								
Chironomidae	25	26	15	7	11	9	93	
Empididae				1			1	
Simuliidae	2	4	1	3	1		11	
Tipulidae	5			4		2	11	
Other	2		2	1	6	2	13	
metrics	A	B	C	D	E	F	Composite	Average
Total Individuals	139	161	102	73	188	87	750	125
Number of Families	15	12	14	14	12	11	25	13
Dominance (1)	0.36	0.44	0.60	0.34	0.51	0.28	0.26	0.42
Dominance (2)	0.54	0.60	0.75	0.45	0.60	0.49	0.44	0.57
Dominant family	Nemou	Taeniop	Nemou	Taeniop	Taeniop	Nemou	Taenioptery	?
Subdominant family	Chirono	Chirono	Chirono	Polycen	Baetida	Ameleti	Nemourida	?

Comparison of Paired Samples

We analyzed matched data collected by EPA and Potesta Associates at 21 sites in Island Creek, Mud River, and Spruce Fork over 3 sampling periods from Summer 1999 to Winter 2000. EPA sampled using its D-frame method described above, and Potesta used the 6-Surber method described above. EPA also took an additional 21 samples using both methods, at 10 different sites. Sample crews visited sites simultaneously. The objective of this analysis was to determine the comparability of samples collected using two different methods. If sample pairs collected in both ways, at the same site and time, show no bias relative to each other, then the two sampling methods would be considered comparable and valid for assessments.

Figure E-1 shows the cumulative number of families in 6 Surbers at 5 representative sites, showing that each successive Surber captures new families not captured by the previous Surbers.

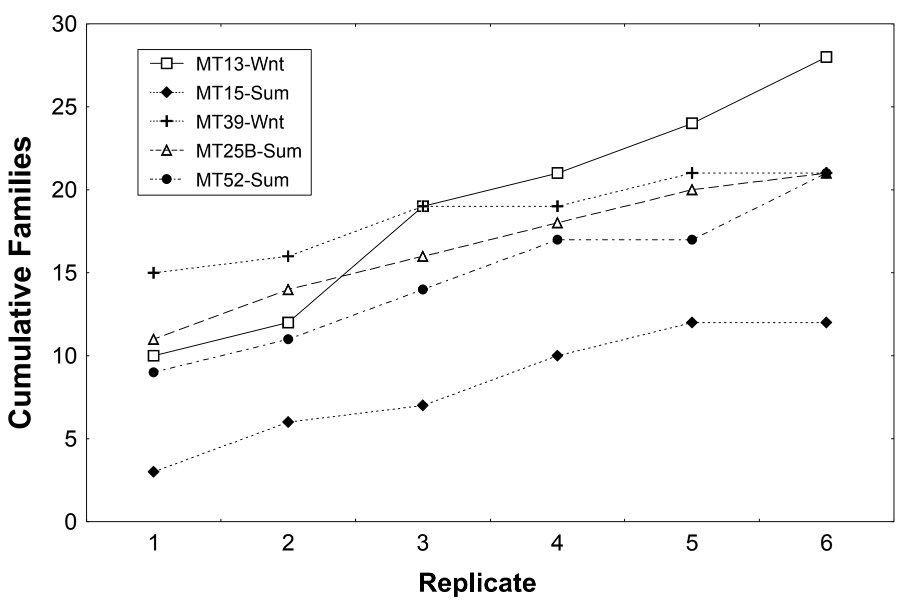


Figure E-1. Cumulative number of families identified in successive Surber samplers from 5 MTM sites.

If we consider the number of organisms captured per unit area of the stream bottom, the 2 methods are unbiased. Figure E-2 compares the individuals per square meter as estimated using Surbers, with individuals per square meter estimated using D-frame samples. The diagonal dotted line represents exact agreement (1:1). While there is scatter about the line, there is no bias above or below the line. Note that Potesta and EPA samples overlap and are unbiased with respect to each other.

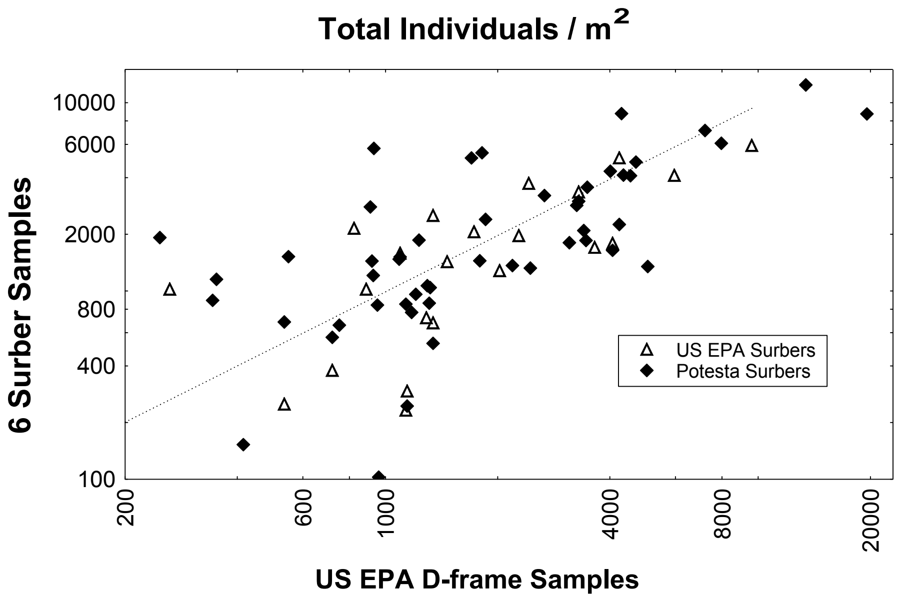


Figure E-2. Total number of individuals from 6 Surber samplers and from EPA D-frame samples. Each point represents a comparison of Surber and D-frame results from the same site at the same time. The vertical axis is the Surber results, and the horizontal axis is the D-frame results. The dotted line is the 1:1 slope of exact agreement between methods. Potesta Surber results are shown with solid diamonds; EPA Surbers with open triangles. All D-frame samples were from EPA.

As explained above, calculating the average number of families from 6 Surbers underestimates richness, since each individual Surber underestimates richness. This is shown graphically in Figure E-3. The average number of families from the Surbers is shown on the vertical axis, and the total families from the D-frame on the horizontal axis. Nearly all the points lie below the 1:1 line. The average bias is approximately 5 families. If we plot the total, cumulative families using Surbers against those using D-frames (Figure E-4), then the D-frames underestimate relative to the Surbers by about 5 taxa, because the D-frames were subsampled to 1/8th the total sample volume. However, if both Surber and D-frame samples are composited and standardized to a constant number of organisms (200), then there is no bias in the family richness (Figure E-5). Note also in Figure 5 that the scatter of points about the 1:1 line is much smaller than for the unstandardized data shown in Figures 3 and 4, and that both Potesta and EPA Surber are unbiased to each other (note 2 symbols in figure).

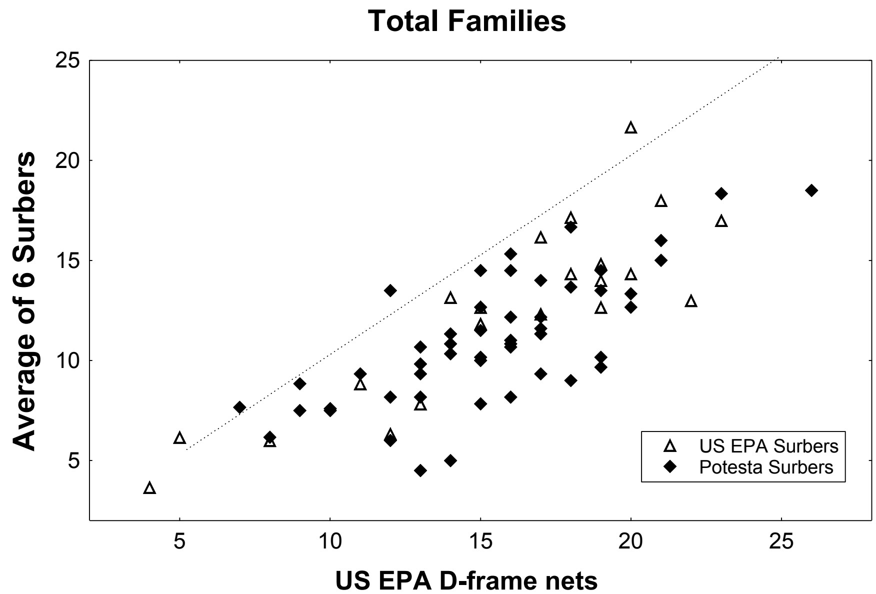


Figure E-3. Number of families per site, averaged over 6 Surbers (vertical), against total numbers from D-frame samples. See Figure 2 caption.

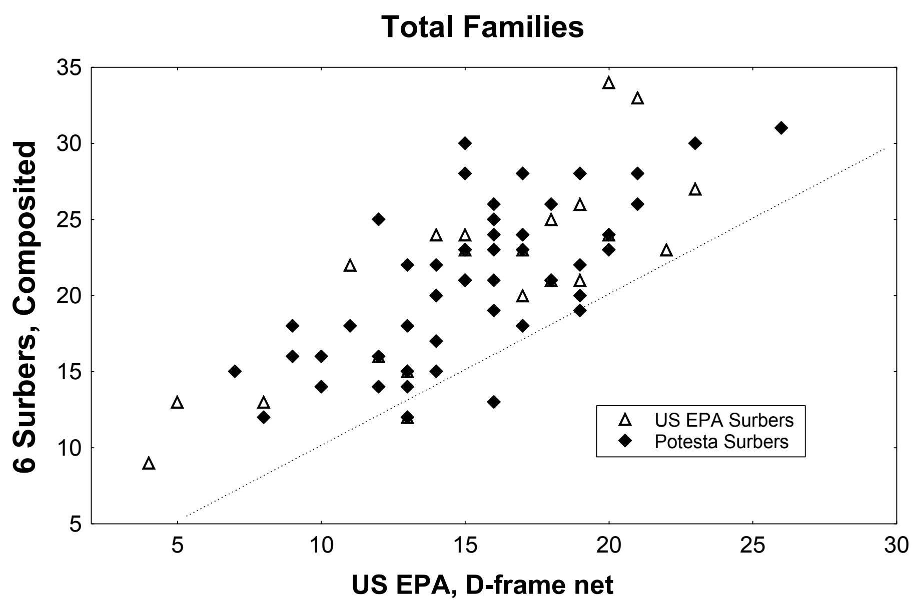


Figure E-4. Total families per site, from composite of 6 Surbers (cumulative), compared to EPA D-frame results. As in Figures 2 and 3.

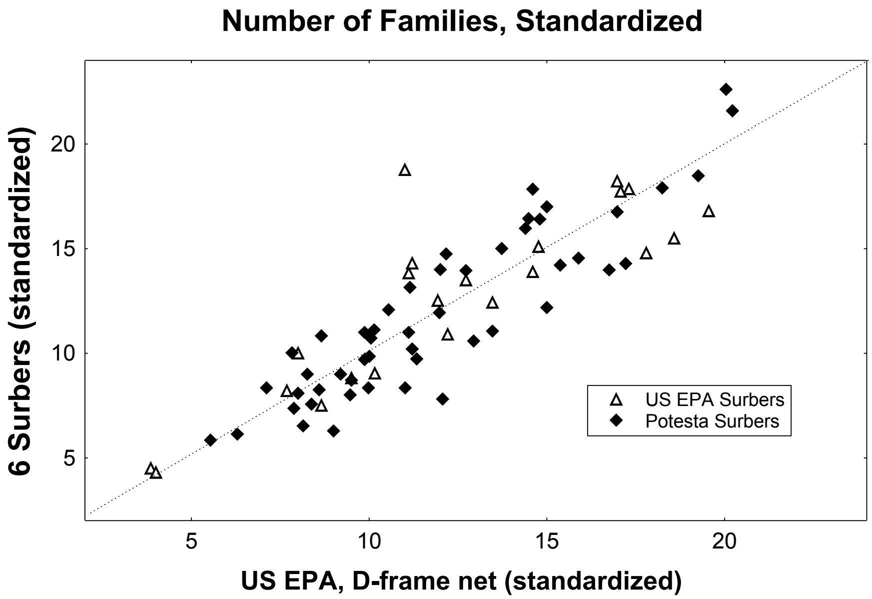


Figure E-5. Number of taxa in standardized Surber samples (vertical) compared to standardized D-frame samples (horizontal). As in Figures 2-4.

The West Virginia Stream Condition Index (WV SCI) is calculated from 6 metric scores. When the index was developed, the scoring formulas were calibrated to a 200 organism sample (Gerritsen et al. 2000). If samples were larger than 200 organisms, they were standardized before the scoring formulas were applied.

Summary: Standardization of Benthic Data

In summary, the data collected by the participants differed in sampling, subsampling and reporting methods. Despite the differences, any one of these sampling, subsampling, and reporting methods is unbiased with respect to the types of organisms collected (all used the same mesh size), the density of organisms (numbers per unit area), and the relative abundances (percent of community). The only bias is that of the number of families (taxa richness) as affected by sampling effort. Sampling effort is a combination of the total area sampled, the heterogeneity of the stream bottom sampled, and the size of the subsample. Since all participants used the same field methods for the D-frame samples, 4 D-frames in the field, use of the D-frame data standardizes the field sampling effort. However, EPA subsampled to 1/8th of the total material (with some exceptions noted in the data); REIC to 1/4th the total material (with some exceptions); and all others counted the entire sample. Therefore, taxa richness was standardized to be equivalent to a subsample of 1/8th the total, original material. Unfortunately, REIC data was reported as combined D-frame and Surber samples and could not be standardized for both sampling effort and subsampling in the laboratory.

Metric Calculations for Benthic Data

The West Virginia Stream Condition Index (WV SCI) rates a site using an average of six standard indices, or metrics, each of which assesses a different aspect of stream health.

The WV SCI metrics include:

- Total Taxa – a count of the total number of families found in the sample. This is a measure of diversity, or richness, and is expected to increase with stream health.
- Number of EPT Taxa – a count of the number of families belonging to the Orders Ephemeroptera (mayflies), Plecoptera (stoneflies), or Tricoptera (caddisflies) Members of these three insect orders tend to be sensitive to pollution. The number tends to increase with stream health.
- Percent EPTs (Number of EPT families / Total number of Families) - this measures the contribution of the pollution-sensitive EPT families to the total benthic macroinvertebrate community. It tends to increase with stream health.
- Percent Chironomidae – the percentage of pollution-tolerant midge (gnat) larvae in the family Chironomidae tends to decrease in healthy streams and increase in streams that are subjected to organic pollution.
- Percent 2 dominant families - a measure of diversity of the stream benthic community. This metric tends to decrease with stream health.
- Hilsenhoff Biotic Index (HBI). The HBI assigns a pollution tolerance value to each family (more pollution-tolerant taxa receive a higher tolerance value). Tolerance values were found in the literature (Hilsenhoff 1987, Barbour et al. 1999) or were assigned by EPA biologists from Wheeling, WV or Cincinnati, OH. The HBI is then calculated by averaging the tolerance values of each specimen in a sample. The HBI tends to increase as water quality decreases

Several taxa were excluded from the analysis because they inhabit terrestrial, marginal, or surface areas of the stream. The excluded taxa included Aranae, Arachnida, Collembola, and Cossidae.

After all the benthic data had been migrated to EDAS, and after all the data had been collapsed to the Family level, the six WV SCI metrics were calculated from composited enumerations, or counts.

Metric Scoring and Index Calculation

As discussed previously, richness metrics are affected by sampling effort, and were therefore standardized to a 100 or 200 organism subsample before scoring. Other WV SCI metrics are independent of sampling effort and did not require standardization. Each of the metrics was then scored on a scale of 0 to 100 using scoring formulae derived for 100 and 200 organism subsamples (Table E-6). The WV SCI was calculated as an average of the six metric scores.

Table E-6. WV SCI: Metric scoring formulas. The richness metrics have two scoring formulas each, depending on the standardized sample size (100 or 200 organisms). The

scoring formulas are from unpublished analyses for 100 organism richness metrics and Gerritsen et al. (2000) for 200 organism richness metrics and other metrics.

Metrics that decrease with stress	***Scoring formulas (X=metric value)***
Total taxa	**$Score_{100} = 100 \times (X/18)$, $Score_{200} = 100 \times (X/21)$**
EPT taxa	**$Score_{100} = 100 \times (X/12)$, $Score_{200} = 100 \times (X/13)$**
% EPT	**score = 100 × (X/91.9)**
Metrics that increase with stress	
%Chironomidae	**score = 100 × [(100-X)/(100-0.98)]**
% 2 dominant	**score = 100 × [(100-X)/(100-36.0)]**
HBI	**score = 100 × [(10-X)/(10-2.9)]**

References

Barbour, M.T., J. Gerritsen, B.D. Snyder, J. B. Stribling. 1999. *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish.* 2nd edition.

Gerritsen, J., J. Burton, M.T. Barbour. 2000. *A stream condition index for West Virginia wadeable streams.*

Vinson, M.R., and C.P. Hawkins. 1996. Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams. Journal of the North American Benthological Society. 15:392-399.

Hilsenhoff, W. L. 1987. *An improved biotic index of organic stream pollution*. Great Lakes Entomologist 20:31-39.

Hurlbert, S.H. 1971. *The nonconcept of Species Diversity: a Critique and Alternative Parameters.* Ecology 52(4): 577-586.

McCormick, F. H., R. M. Hughes, P. R. Kaufmann, D. V. Peck, J. L. Stoddard, A. T. Herlihy. 2001. Development of an index of biotic integrity for the Mid-Atlantic Highlands region. Transactions of the American Fisheries Society 130:857-877.